Study on the differentially expressed genes and signaling pathways in dermatomyositis using integrated bioinformatics method

Wei Liu, MD\textsuperscript{a,b,*}, Wen-Jia Zhao, MD\textsuperscript{a}, Yuan-Hao Wu, PhD\textsuperscript{a,b}

Abstract

Dermatomyositis is a common connective tissue disease. The occurrence and development of dermatomyositis is a result of multiple factors, but its exact pathogenesis has not been fully elucidated. Here, we used biological information method to explore and predict the major disease related genes of dermatomyositis and to find the underlying pathogenic molecular mechanism.

The gene expression data of GDS1956, GDS2153, GDS2855, and GDS3417 including 94 specimens, 66 cases of dermatomyositis specimens and 28 cases of normal specimens, were obtained from the Gene Expression Omnibus database. The 4 microarray gene data groups were combined to get differentially expressed genes (DEGs). The gene ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichments of DEGs were operated by the database for annotation, visualization and integrated discovery and KEGG orthology based annotation system databases, separately. The protein–protein interaction networks of the DEGs were built from the STRING website. A total of 4097 DEGs were extracted from the 4 Gene Expression Omnibus datasets, of which 2213 genes were upregulated, and 1884 genes were downregulated. Gene ontology analysis indicated that the biological functions of DEGs focused primarily on response to virus, type I interferon signaling pathway and negative regulation of viral genome replication. The main cellular components include extracellular space, cytoplasm, and blood microparticle. The molecular functions include protein binding, double-stranded RNA binding and MHC class I protein binding. KEGG pathway analysis showed that these DEGs were mainly involved in the toll-like receptor signaling pathway, cytosolic DNA-sensing pathway, RIG-I-like receptor signaling pathway, complement and coagulation cascades, arginine and proline metabolism, phagosome signaling pathway. The following 13 closely related genes, XAF1, NT5E, UGCG, GBP2, TLR3, DDX58, STAT1, GBP1, PLSCR1, OAS3, SP100, IGK, and RSAD2, were key nodes from the protein–protein interaction network.

This research suggests that exploring for DEGs and pathways in dermatomyositis using integrated bioinformatics methods could help us realize the molecular mechanism underlying the development of dermatomyositis, be of actual implication for the early detection and prophylaxis of dermatomyositis and afford reliable goals for the curing of dermatomyositis.

Abbreviations: DAVID = the database for annotation, visualization and integrated discovery, DEGs = differentially expressed genes, GEO = Gene Expression Omnibus, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, KOBAS = KEGG orthology based annotation system, PPI = protein–protein interaction, RRA = robust rank aggregation.

Keywords: dermatomyositis, differentially expressed genes, Gene Expression Omnibus data, integrated bioinformatics

1. Introduction

Dermatomyositis is a non-suppurative inflammatory disease mainly involving rhabdomyositis and lymphocyte infiltration.\cite{1} It can be accompanied or not accompanied by multiple skin lesions.\cite{2} This disease is characterized by symmetrical weakness of limb band, neck and pharynx muscles, often involving multiple organs, and may be associated with tumors and other connective tissue diseases.\cite{3} The exact cause of the disease is not yet clear, and it is generally believed to be related to heredity and viral infection.\cite{4} The incidence of polymyositis and dermatomyositis has significant racial differences.\cite{5} African Americans had the highest incidence, with the proportion of blacks to whites ranging...
from 3 to 4:1. The incidence of dermatomyositis in children is higher in Asia and Africa than in Europe and America. At present, the diagnosis of dermatomyositis is mainly based on symmetrical proximal muscle weakness, pain and tenderness, accompanied by characteristic skin lesions such as periorbital purple-red edema, Gottron sign and nail fold rigid dilated capillary erythema. Increased muscle enzyme activity is one of the important serological indicators for the diagnosis of this disease. The serum myoglobin level in most patients with myositis increased and was parallel to the condition of the disease. Patients with dermatomyositis may develop autoantibodies such as anti-nuclear antibodies and anti-Jo-1 antibodies. However, the positive rate of diagnosis is low and early diagnosis is difficult. Glucocorticoid, methotrexate, azathioprine, and other immunosuppressive agents are commonly used medicines for this disease. For patients with dermatomyositis complicated by heart or lung disease, the condition is often serious, and the treatment effect is poor. In recent years, although there are some new immunotherapy programs for dermatomyositis, its efficacy still needs to be improved. Therefore, it is very important for the prevention and treatment of dermatomyositis to find new early diagnostic indicators and explore possible therapeutic targets.

The dynamic expression of tens of thousands of genes or gene transcripts can be detected on a large scale by DNA microarray technology. In recent years, DNA microarray technology has developed rapidly. By comparing gene expression differences between normal tissue cells and pathological tissue cells, a subset of characteristic genes closely related to disease types can be obtained. By analyzing the differential expression of genes in different pathological tissues and observing the correlation between genes, we can find the mechanism of pathological changes from the differential expression of genes. However, due to the high dimension, small sample size and non-linear distribution of gene expression profile data, direct data analysis often results in “dimension disaster” and “over-fitting” phenomena, which make the research progress slow, and the results are not ideal or even cannot be used. Therefore, it is necessary to develop an efficient feature gene selection method to process the data and select a gene subset that is of guiding significance for pathological changes, to facilitate diagnosis.

Robust rank aggregation (RRA) method is an algorithm using probability model aggregation. It was proposed by KOLDE in 2012. For ordered gene list, this algorithm can make good use of the influence of parameters, outlier reduction, and noise of the Gen model optimization algorithm on data results, leaving only the genes related to statistics in the final gene list. In research practice, many bioinformatics analysis tools can output gene lists sorted by some sort for us. When using different analysis methods or different data sets for analysis, many different gene lists can be output. RRA method can be used to correlate these different gene lists and obtain a new gene list that eliminates interference.

In this study, RRA was used to analyze the differentially expressed genes (DEGs) of dermatomyositis, hoping to find out the DEGs of dermatomyositis, and provide new ideas for the early diagnosis and gene therapy of dermatomyositis.

First of all, we downloaded 4 original microarray datasets, GDS1956, GDS2153, GDS2855, and GDS3417 were downloaded using the keyword “dermatomyositis” searched in the GEO database. GDS1956 is a GPL96 platform, [HG-U133A] Affymetrix Human Genome U133A Array, including 10 samples of normal human skeletal muscle subsets and 21 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS2153 is GPL96, [HG-U133A] Affymetrix Human Genome U133A Array, which contains 4 samples of normal human skeletal muscle subsets and 3 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS2855 is GPL97, [HG-U133A] Affymetrix Human Genome U133B Array, which consists of 10 samples of normal human skeletal muscle subsets and 21 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS3417 is GPL96, [HG-U133A] Affymetrix Human Genome U133A Array, which contains 4 samples of normal human skeletal muscle subsets and 19 samples of skeletal muscle subsets of dermatomyositis patients. The platform and series matrix files are downloaded as TXT files. Inter chip standardization can determine a reference chip, assume some invariants, and calibrate the data of other chips as a whole. The microarray data were processed by the expresso function in affy software package in 3 steps: background correction, standardization and summary, and the standardized gene expression matrix was obtained.

2. Methods

GEO database preserves many gene expression data generated by the scientific community. Researchers can observe a great deal of evidence to show the behavior of a gene or a group of genes. Users can mine different genes to obtain evidence to confirm laboratory findings and can also search for candidate genes in GEO database that deserve further research in the laboratory. Combining sequence information with expression information is helpful for functional annotation and characterization of unknown genes, or for discovering new roles of characteristic genes. These data are also valuable for genome-wide research, enabling biologists to examine global gene expression in different cell types and states, compare homologous genes with other species, and find repetitive patterns that help to form hypothetical co-regulatory transcripts in functional networks and pathways. Ethical approval was not necessary in our study because the expression profiles were downloaded from the public database and no new experiments in patients or animals were performed.

The gene expression profiles of GDS1956, GDS2153, GDS2855, and GDS3417 were downloaded using the keyword “dermatomyositis” searched in the GEO database. GDS1956 is a GPL96 platform, [HG-U133A] Affymetrix Human Genome U133A Array, including 10 samples of normal human skeletal muscle subsets and 21 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS2153 is GPL96, [HG-U133A] Affymetrix Human Genome U133A Array, which contains 4 samples of normal human skeletal muscle subsets and 3 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS2855 is GPL97, [HG-U133B] Affymetrix Human Genome U133B Array, which consists of 10 samples of normal human skeletal muscle subsets and 21 samples of skeletal muscle subsets of dermatomyositis patients. The platform of GDS3417 is GPL96, [HG-U133A] Affymetrix Human Genome U133A Array, which contains 4 samples of normal human skeletal muscle subsets and 19 samples of skeletal muscle subsets of dermatomyositis patients. The platform and series matrix files are downloaded as TXT files. Inter chip standardization can determine a reference chip, assume some invariants, and calibrate the data of other chips as a whole. The microarray data were processed by the expresso function in affy software package in 3 steps: background correction, standardization and summary, and the standardized gene expression matrix was obtained.
2.1. Screening for DEGs

The purpose of screening DEGs is to find genes with significant changes in different experimental samples. It is of great significance to find molecular markers, therapeutic targets, and molecular biological mechanisms that are helpful for disease diagnosis. The obtained DEG platform and matrix file series are converted by using R language software and annotation package. The ID corresponding to probe name is transformed to the international canonical name of gene (gene symbol) and stored in TXT file. Gene diverse expression analysis was made using Limma software package in Bioconductor Package (available online: http://www.bioconductor.org/).[34]

The correlative manipulating instruction codes were entered into R, and the DEGs in dermatomyositis and normal specimens of the 4 microarray datasets were calculated by the limma software package. Specimens with an adjusted P-value of <.05 and log fold change (FC) >2 were treated as DEGs. The TXT outcomes were saved for later calculation.

2.2. Integration of microarray data

The list of DEGs from the 4 microarray datasets calculated by limma packet analysis was stored as a TXT file. The RRA software package was downloaded, and R was applied to run the command code. A portion of genes that were up- or down-regulated in the 4 chips were saved for later research. The RRA approach is openly workable in the Comprehensive R Network (http://cran.r-project.org/).[33]

2.3. GO and KEGG pathway enrichment analyses of DEGs

GO ontology analysis is to annotate and classify genes by biological process (BP), molecular function (MF), and cellular component. By GO analysis, not only the functions of differential genes can be identified, but also the functions of DEGs can be studied. Through the statistical analysis of the function enrichment degree of DEGs, we can calculate the p-value of the corresponding function of DEGs and the FDR value of P-value, that is, Q-value. Finally, combined with the research purpose, we can locate the most likely function related to DEGs. The analysis of KEGG pathway enrichment is based on KEGG pathway. Using fish statistical test, we can find out the signal pathway of significant enrichment of DEGs, which helps us to further clarify the biological function of genes, understand which metabolic pathway genes are enriched in, and screen out the key signal pathway.

The DAVID database is a bioinformatics database, which integrates biological data and analysis tools, provides systematic and comprehensive bioinformatics annotations for large-scale gene or protein lists, and identifies the most significant bioinformatics annotations by statistical methods. GO enrichment analysis were carried out by a DAVID online tool for tracing DEGs. KEGG pathway annotation of DEGs was carried out by the KOBAS online computing database. We logged in the STRING and KOBAS online computing database, inputted the gene name of the DEGs, screened out the path with high enrichment degree according to the threshold value of “P < .05”, and then used R software for visualization.

2.4. PPI network integration

The STRING database (http://string-db.org/) is a software system for searching for protein-protein interactions, which helps to mine the core regulatory genes. It includes not only the direct physical interaction between proteins, but also the indirect functional correlation between proteins. Each node in the network graph represents a protein, and the lines between nodes represent the interaction between 2 proteins. Different colors correspond to different interaction types.[35] All protein-protein interaction data are weighted, integrated, and there will be a calculated reliable value, and all calculated predictions can be viewed through a specific view. The corresponding proteins in the central node may be core proteins or key candidate genes with important physiological regulatory functions.[36] The DEGs were input into the string database to construct the interaction network of differential gene coding proteins. By analyzing the network structure, the local subnetworks formed in the whole network and the key node proteins were obtained.

3. Results

3.1. Microarray data information and identification of DEGs in dermatomyositis

The dermatomyositis expression microarray datasets GDS1956, GDS2153, GDS2855, and GDS3417 were standardized, and the results are shown in Figure 1. When the GDS1956 dataset was screened by the limma package (corrected P-value <.05, logFC > 2), 681 DEGs were obtained. Among them, 468 upregulated genes and 195 downregulated gene were identified. 1362 DEGs were screened from the GDS2153 dataset, including 733 upregulated genes and 629 downregulated genes. 900 DEGs were screened from the GDS2855 dataset, including 600 upregulated genes and 300 downregulated genes. Additionally, 2675 DEGs were screened from the GDS3417 dataset, including 1376 upregulated genes and 1299 downregulated genes. The differential expression of multiple genes from the 2 sets of sample data included in each of the 4 microarrays is shown in Figure 2. The cluster heatmaps of the top 20 DEGs are shown in Figure 3.

DEGs in dermatomyositis was identified by comprehensive bioinformatics. The microarray data of 4 kinds of dermatomyositis gene expression were analyzed by limma software package. The data were classified according to logarithmic variation, and then by RRA analysis was performed (corrected P-value <.05). The RRA method assumes that each gene in each experiment is randomly arranged. If a gene ranks higher in all experiments, the smaller its P-value, the greater the possibility of differential gene expression. Using R-heatmap software, the Heatmap of Top 20 up- and downregulated genes of each dataset was drawn, as shown in Figure 3.

3.2. GO term enrichment analysis of DEGs

Biological annotation of the DEGs in dermatomyositis identified from an integrated analysis of microarray data was performed using the DAVID online analysis tool and GO functional enrichments of up- and downregulated genes with a P-value of <.05 were obtained. GO analysis of DEGs was divided into 3 functional groups, including MF, biological processes, and cell composition.[38] The results are shown in Figures 4 and 5. Significant results of the GO enrichment analysis of DEGs in dermatomyositis are shown in Table 1. The biological functions of DEGs focused primarily on response to virus, type I interferon signaling pathway and negative regulation of viral genome replication. The main cellular components include extracellular...
space, cytoplasm and blood microparticle. The MFs include protein binding, double-stranded RNA binding and MHC class I protein binding. These results indicate that most DEGs are significantly enriched in cell proliferation, binding, cell cycle regulation, and transcriptional activity.

3.3. **KEGG pathway analysis of DEGs**

Using the KOBAS online analysis database (http://kobas.cbi.pku.edu.cn/) to analyze the DEGs identified from dermatomyositis integrated gene microarrays, the most significantly enriched pathways of the DEGs were submitted to KEGG analysis. The

---

*Figure 1.* Standardization of gene expression. (A) The standardization of GDS1956 data, (B) the standardization of GDS2153 data, (C) the standardization of GDS2855 data, (D) the standardization of GDS3417 data. The green bar represents the data before normalization, and the red bar represents the normalized data.
results are shown in Table 2. The signaling pathways of DEGs were mainly enriched in the toll-like receptor signaling pathway, cytosolic DNA-sensing pathway, RIG-I-like receptor signaling pathway, complement and coagulation cascades, arginine and proline metabolism, phagosome signaling pathway in dermatomyositis. The data were imported into Cytoscape to calculate the topological characteristics of the network and determine each node.

3.4. Analysis of common genes in 4 microarray datasets

The number of unique and common genes in 4 kinds of microarray datasets was analyzed by Venn diagram,[39] and the similarity and overlap of gene composition among different microarray samples were further compared. The results showed that the number of identical genes in the 4 microarray datasets was 13, as shown in Figure 6. The common genes are XAF1, NT5E, UGCG, GBP2, TLR3, DDX58, STAT1, GBP1, PLSCR1, OAS3, SP100, IGK, and RSAD2.

3.5. Analyzing DEGs in dermatomyositis using a PPI network

13 genes are common to 4 microarray datasets. The common expression products in dermatomyositis were constructed using the STRING database to construct PPI networks. After removing the isolated and partially connected nodes, a complex network of DEGs was constructed, as shown in Figure 7.
4. Discussion

We built a PPI network of protein encoded by DEGs and screened out the following 13 genes that have distinct relation with dermatomyositis: XAF1, NT5E, UGCG, GBP2, TLR3, DDX58, STAT1, GBP1, PLSCR1, OAS3, SP100, IGK, and RSAD2. The proteins expressed by these genes are critical nodes in the PPI network. Pathway enrichment analysis pointed out that the genes mostly participated in the toll-like receptor signaling pathway, cytosolic DNA-sensing pathway, RIG-I-like receptor signaling pathway, complement and coagulation cascades, arginine and proline metabolism, phagosome signaling pathway. X chromosome linked inhibitor of apoptosis protein inhibits apoptosis by binding to caspases and inhibiting the function of caspases. XAF1 is widely expressed in all embryonic and adult normal tissues, but it is low in cancer cell lines and tumor tissues. It was found that apoptotic nuclei could be seen in the muscle tissue of dermatomyositis, and the expression of apoptotic factors was upregulated, suggesting that apoptotic cells might be involved in the pathological process of dermatomyositis. XAF1-related excessive apoptosis may exist in dermatomyositis, which may be one of the causes of dermatomyositis.

NT5E gene is located on chromosome 6 q14-q21, and its expression product NT5E is a membrane nuclease anchored by GPI with a size of about 69 kDa. NT5E not only hydrolyzes adenine as a nucleotidase, but also has many physiological
functions. Because the activity of NT5E in peripheral blood T cells is 10 times higher than that in thymus, and the activity of NT5E in peripheral B cells is 5 to 6 times higher than that in fetal spleen or umbilical cord blood, it is a marker of lymphocyte maturation. In addition, NT5E has cell-specific function during neuronal development and plasticity. Ring et al. found that CD4+ CD25+ Treg cells can downregulate the expression of Bcl-2 protein to activate NF-κB to promote cell death. Anti-TNF-α monoclonal antibody can inhibit the production of TNF-α, which can mediate cellular active immunity (natural immunity). TLR3 specifically recognizes ds-RNA, the intermediate product of viral replication, to activate NF-κB and IFN-β precursors. Doyle et al. confirmed that anti-TLR3 monoclonal antibody can inhibit the production of IFN-β in fibroblasts. TLR3 pathway can enhance the expression of inflammatory cytokines such as IL-1β and IL-6, amplify the local inflammatory response of patients with dermatomyositis.

Table 1

| Category               | GO ID       | Term                                      | P-value       | Genes                                                                 |
|------------------------|-------------|--------------------------------------------|---------------|-----------------------------------------------------------------------|
| Biological process     | GO:0060337  | Type I interferon signaling pathway        | 5.0429E-32    | SP100, BST2, IFIT2, IFIT3, OAS3, RSAD2, OAS1, HLA-C, OAS2, HLA-E, STAT1, PSMB8, SSO20, IFT3, IFT2, IG15, IFIT2, IFIT2, XAF1, MX1, MX2, IFN6, ADAR |
| Biological process     | GO:0051607  | Defense response to virus                  | 4.5689E-31    | ZC3H11V, IFIT2, IFIT3, OAS3, TLR3, RSAD2, OAS1, IFH4L, OAS2, PMAIP1, SSO20, IG15, DDX60, MX1, APOBEC3A_B, MX2, PTPRC, BST2, HERC5, IFIT1, STA1, TRM22, IFIT3, PSRC1, IFIT3, IFT2, EGF2K2, ADAR |
| Biological process     | GO:0009615  | Response to virus                          | 3.6477E-26    | IFIT1, ZC3H11V, BST2, IFIT2, IFIT3, OAS3, CCL19, RSAD2, OAS1, IFIT2, IFIT3, DDX58, IFT3, DDX60, IF7, DDX21, EGF2K2, MX1, MX2, ADAR |
| Biological process     | GO:0045071  | Negative regulation of viral genome replication | 8.1474E-21    | BST2, ZC3H11V, IFIT2, IFIT3, OAS3, RSAD2, OAS1, IFIT6, SSO20, PSRC1, IG15, EGF2K2, APOBEC3A_B, MX1, ADAR |
| Biological process     | GO:0045087  | Innate immune response                     | 3.2882E-16    | IFIT1, HMGB2, ZC3H11V, TLR2, TLR3, C1R, C1S, NOD2, DDX60, IGH2, APOBEC3A_B, MX1, MX2, BST2, LY96, HERC5, HLA-C, IFIT1, HLA-E, DDX58, CIQB, IFIT5, EGF2K2, IGL1, ADAR |
| Biological process     | GO:0060333  | Interferon-gamma-mediated signaling pathway | 9.0838E-14    | TRM38, NMI, S100, IFIT7, OAS3, OAS1, HLA-C, OAS2, STA1, HLA-E, TRM22 |
| Biological process     | GO:0006955  | Immune response                            | 1.4649E-10    | IGLV1-44, IFIT2, IL7, IL18, CCR1, IFITM3, OAS3, TLR2, CCL19, OAS1, HLA-C, C1R, OAS2, HLA-E, TRM22, LILRB2, IG2A, IGL1, IFIT6 |
| Biological process     | GO:0035456  | Response to interferon-beta                | 2.1107E-09    | PSRC1, BST2, IFITM2, IFITM3, XAF1, STA1 |
| Biological process     | GO:0035455  | Response to interferon-alpha               | 4.1973E-09    | BST2, BCL2, IFITM2, IFITM3, XAF1, STA1 |
| Molecular function     | GO:0003725  | Double-stranded RNA binding                | 9.0650E-09    | DDX58, IFIT1, DDX60, OAS3, TLR3, DDX21, OAS1, OAS2, EGF2K2 |
| Biological process     | GO:0032728  | Positive regulation of interferon-beta production | 3.0748E-08    | DDX58, IFIT1, HMGB2, ZC3H11V, IFIT7, TLR2, TLR3 |
| Biological process     | GO:0032760  | Positive regulation of tumor necrosis factor production | 2.0683E-05    | NOD2, LIF6, TLR2, TLR3, CCL19, HLA-E |
| Biological process     | GO:0034341  | Response to interferon-gamma               | 2.3596E-05    | SP100, BST2, IFIT2, IFITM2, IFITM3, OAS3, RSAD2, TLR3, OAS1, HLA-C, IFIT2, DDX58, IFIT3, DDX60, IF7 |
| Biological process     | GO:0032727  | Positive regulation of interferon-alpha production | 7.4704E-05    | BST2, IFIT2, IFITM2, IFITM3, XAF1, STA1 |
| Biological process     | GO:0006952  | Defense response                           | 0.00011621    | NOD2, TAP1, RNASE6, HCP5, MX1, MX2 |
| Molecular function     | GO:0001730  | 2’-5’-oligoadenylate synthetase activity   | 0.00028316    | OAS2, OAS1, OAS2 |
| Biological process     | GO:0032755  | Positive regulation of interferon-6 production | 0.00029471    | DDX58, LILRB2, NOD2, TLR2, TLR3 |
| Biological process     | GO:0002346  | Positive regulation of NF-kappaB import into nucleus | 0.00430879    | IL18, TLR2, TLR3, CCL19 |
| Biological process     | GO:0009597  | Detection of virus                         | 0.00499383    | DDX58, IFIT1, TLR3 |
| Cellular component     | GO:0030670  | phagocytic vesicle membrane                | 0.00069586    | TLRG1, TLR2, HLA-C, HLA-E, RAB20 |

DEGs = differentially expressed genes, GO = gene ontology.

UDP-glucosylceramide glucosyltransferase (UGCG) can reduce ceramide levels in vivo and induce cells to escape ceramide-induced apoptosis. UGCG gene can regulate apoptosis-related pathways. Increased expression of UGCG can induce increased expression of anti-apoptotic gene Bcl-2, decreased expression of Bax and Caspase-3. Bcl-2 protein blocks cell necrosis by maintaining mitochondrial membrane potential, which can delay or prevent Fas induced myocyte necrosis in patients with dermatomyositis. UGCG and caspase-3 can cause the cleavage of PARP and apoptosis of muscle cells, leading to the occurrence of dermatomyositis.

Guanylate-binding protein 2 (GBP2) is a kind of large GTPase protein with molecular weight of 65–67 kDa induced by interferon (IFN), which can mediate cellular active immunity and participate in many biological reactions such as anti-pathogenic microorganisms. GBP2 can increase the secretion of inflammatory factors, which has a strong immune promoting effect in the pathogenesis of dermatomyositis.

Toll-like receptors (TLR) are important proteins involved in non-specific immunity (natural immunity). TLR3 specifically recognizes ds-RNA, the intermediate product of viral replication, to activate NF-κB and IFN-β precursors. Doyle et al. confirmed that anti-TLR3 monoclonal antibody can inhibit the production of IFN-β in fibroblasts. TLR3 pathway can enhance the expression of inflammatory cytokines such as IL-1β and IL-6, amplify the local inflammatory response of patients with
Figure 4. GO enrichment analysis of DEGs in dermatomyositis. (A) GO analysis divided DEGs into 3 functional groups: molecular function, biological processes, and cell composition. (B) GO enrichment significance items of DEGs in different functional groups. DEGs = differentially expressed genes, GO = gene ontology.

Figure 5. Distribution of DEGs in dermatomyositis for different GO-enriched functions. The color of the stripe that each gene points to indicates the enriched gene pathway. The color depth of each gene block represents the amount of logFC. DEGs = differentially expressed genes, GO = gene ontology.
dermatomyositis, and aggravate the disease. TLR2, TLR4, and TLR9 are elevated in muscle tissue of patients with dermatomyositis, and can activate TNFα, promote the transcription of inflammatory factors.

The DDX58 gene, also known as RIG-I, recognizes viral RNA and activates innate antiviral responses by inducing type I interferon and inflammatory cytokines. DDX58 gene plays an important role in activating NF-κB pathway and plays an

| KEGG ID | Term                                         | P-value  | FDR        | Genes                                                                 |
|---------|----------------------------------------------|----------|------------|----------------------------------------------------------------------|
| hsa05168 | Herpes simplex infection                     | 3.28E-10 | 3.56E-07   | IFIH1, SP100, OAS3, TLR2, TLR3, HLA-C, OAS1, OAS2, HLA-E, STAT1, DDX58, IFN7, TLR1, TLR2 |
| hsa05164 | Influenza A                                  | 2.47E-00 | 2.68E-06   | DDX58, IFIH1, IL18, IRF7, OAS1, TLR3, RSAD2, OAS1, OAS2, EF2AK2, MX1, STAT1, ADAR |
| hsa05162 | Measles                                      | 2.77E-08 | 3.00E-05   | DDX58, IFIH1, IFN7, OAS3, TLR2, OAS1, OAS2, EF2AK2, MX1, STAT1, ADAR |
| hsa05160 | Hepatitis C                                  | 4.31E-05 | 3.56E-07   | DDX58, IFN7, OAS1, IFN7, TLR2, TLR3, RSAD2, OAS2, EF2AK2, STAT1 |
| hsa04150 | Phagosome                                    | 0.003837786 | 4.088134319 | DDX58, IFN7, TLR2, TLR3, STAT1                                      |
| hsa00330 | Arginine and proline metabolism              | 0.005271044 | 5.8421963   | SAT1, LAP3, GADD32, GMAT                                           |
| hsa04620 | Toll-like receptor signaling pathway         | 0.00712164 | 7.46084832  | LY96, IFN7, TLR2, TLR3, STAT1                                      |
| hsa04623 | Cytosolic DNA-sensing pathway                | 0.011030567 | 11.3490913  | DDX58, IFN7, STAT1                                                 |
| hsa05321 | Inflammatory bowel disease (IBD)             | 0.011030567 | 11.3490913  | NOD2, IL18, TLR2, STAT1                                            |
| hsa04622 | RIG-I-like receptor signaling pathway        | 0.014059114 | 14.24698402 | DDX58, IFN7, ISG15, IFN7                                           |
| hsa05203 | Viral carcinogenesis                          | 0.015985325 | 16.04820622 | SP100, IFN7, HLA-C, PMAAP1, EF2AK2, HLA-E                          |
| hsa05133 | Pertussis                                    | 0.016910157 | 16.9076255  | C1QB, HLA-C, CR1, C1S                                             |
| hsa05152 | Tuberculosis                                 | 0.03897259 | 35.0481367  | C1QB, NOD2, IL18, TLR2, STAT1                                     |
| hsa05150 | Staphylococcus aureus infection              | 0.058318374 | 47.91999402 | DDX58, HLA-C, EF2AK2, HLA-E                                        |
| hsa05169 | Epstein-Barr virus infection                 | 0.058411269 | 47.96974322 | DDX58, HLA-C, EF2AK2, HLA-E                                        |
| hsa04610 | Complement and coagulation cascades          | 0.089332034 | 63.78894865 | C1QB, HLA-C, C1S                                                  |

DEGs = differentially expressed genes, KEGG = Kyoto Encyclopedia of Genes and Genomes.

Figure 6. Venn diagram of 4 microarray datasets. Each color represents 1 expression microarray dataset. There are 13 common genes in the intersection of the 4 expression microarray datasets.
important role in the pathogenesis of dermatomyositis.\textsuperscript{[60]} It has been found that \textit{DDX58} gene is significantly up regulated in psoriasis patients and psoriasis like mice.\textsuperscript{[61]}

Signal transducer and activator of transcription 1 (STAT1) is involved in many basic biological processes, such as cell proliferation, apoptosis, angiogenesis, and immune response.\textsuperscript{[62]} Type I interferon can participate in the myopathy of patients with dermatomyositis through STAT1 signaling pathway.\textsuperscript{[63]} IFN\textsubscript{a} and IFN\textsubscript{b} can promote the expression of type I interferon induced gene by binding with IFNAR1 and IFNAR2 to activate STAT1 signaling pathway, leading to muscle tissue damage.\textsuperscript{[64]} Recently, it has been confirmed that STAT1 expression in the muscle tissue of patients with dermatomyositis is significantly higher than that of the normal control group.\textsuperscript{[65]}

Guanylate binding protein 1 (GBP1) is a kind of monomer engine protein composed of 593 amino acids.\textsuperscript{[66]} There is no expression of GBP1 in normal human skin, but when the skin is invaded by inflammation, the expression of GBP1 is significantly increased and the aggravation of inflammatory diseases is inhibited, GBP1 may be a potential early-warning signal of inflammation.\textsuperscript{[67]} In endothelial cells, GBP1 can be strongly induced by IFN-\gamma, IFN-\alpha, IFN-\beta, TNF-\alpha, ILs, and other inflammatory factors, and inhibit the proliferation and invasion of endothelial cells in inflammatory diseases.\textsuperscript{[68]}

Phospholipid scramblase 1 (PLSCR1) has many functions including signal transduction, pathogen invasion/absorption, intracellular transport, secretion, and endocytosis. PLSCR1 can promote the activation of Caspase-3 and protein kinase C and increase the apoptosis induced by etoposide.\textsuperscript{[69]} The main pathological reaction in the course of dermatomyositis is not only degeneration and necrosis of muscle tissue, but also closely related to apoptosis.\textsuperscript{[69]} PLSCR1 induced apoptosis plays an important role in dermatomyositis.\textsuperscript{[70]}

Oligoadenylate synthetase 3 (OAS3) is an important antiviral protein, which is induced by IFN, and plays a key role in regulating protein synthesis and immune response.\textsuperscript{[71]} It has been confirmed that OAS3 is related to Sjogren syndrome complicated with hepatitis B virus infection.\textsuperscript{[72]} At present, there are few studies on the relationship between OAS3 and dermatomyositis, which need further experimental verification.

Anti-solid acid nuclear protein of 100,000 antibodies (SP100) is one of the antinuclear antibodies.\textsuperscript{[73]} SP100 has been confirmed to be involved in viral infection, self-ubiquitination regulation, and play a role in interferon, p53, and other signaling pathways.\textsuperscript{[74]} SP100 is clinically associated with autoimmune hepatitis.\textsuperscript{[75]} Clinically, it often occurs before liver injury in patients with dermatomyositis.\textsuperscript{[76]}

Immunoglobulin \kappa light chain (IGK) gene is a kind of highly conserved housekeeping gene, which is produced and secreted only by plasma cells under normal conditions.\textsuperscript{[77]} It is highly expressed in some immune related diseases such as multiple myeloma, leukemia, lymphoma, autoimmune diseases, and infectious diseases.\textsuperscript{[78]} At present, there are few studies on the role of IGK in the pathogenesis of dermatomyositis. It is suggested that the overexpression of IGK may be involved in the pathogenesis of dermatomyositis by activating NF-\kappaB pathway.

\textbf{Figure 7.} PPI network. Circles represent genes, lines represent the interaction of proteins between genes, and the results within the circle represent the structure of proteins. Line color represents evidence of the interaction between the proteins. PPI = protein–protein interaction.
Radical S-adenosyl methionine domain-containing protein 2 (RSAD2) is an endoplasmic reticulum associated virus suppressor gene mediated by interferon I. RSAD2 promotes Th1 cells to differentiate into Th2 cells, and then induces and regulates B lymphocyte activation and autoantibody secretion. RSAD2 has been found to be related to a variety of autoimmune diseases, such as rheumatoid arthritis, systemic lupus erythematosus, ankylosing spondylitis, and so on.

5. Conclusions

We retrieved several microarray datasets on the GEO website and got 4 microarray datasets. Then, we applied R software and statistics to merge these datasets. We have recognized 4097 candidate DEGs, which may participate in the pathogenic process of dermatomyositis. By enriching the GO and KEGG pathways, we got that DEGs were primarily enriched in the toll-like receptor signaling pathway, cytokine DNA-sensing pathway, RIG-I-like receptor signaling pathway, complement and coagulation cascades, arginine and proline metabolism, phagosomal signaling pathway in dermatomyositis, which offer a foundation for exploring the pathogenic processes of dermatomyositis. We ultimately formed a PPI network of DEGs in dermatomyositis and searched many key genes encoding proteins in the network that participate in the pathological mechanism of dermatomyositis in the form of molecular populations. Following research of this network would be instructive for cognition the relationship between DEGs. These researches ameliorate our cognition of the molecular mechanism of dermatomyositis and the appearance and development of the underlying pathogenic factors. Our study has considerable clinical meaning for the early diagnosis and remedy, in addition to the warding off, of dermatomyositis and obtains available targets for the medication of dermatomyositis. Nevertheless, further in vivo and in vitro verification experiments are demanded to demonstrate the effect of the recognized genes bound up with dermatomyositis.

Author contributions

Wei Liu developed the concept and algorithm. Yuan-Hao Wu and Wen-Jia Zhao integrated the databases, implemented the codes, and established the webservice. Wei Liu, Yuan-Hao Wu, and Wen-Jia Zhao analyzed and evaluated the results. Yuan-Hao Wu and Wen-Jia Zhao drafted the manuscript. All authors have read and approved the final version of the manuscript for publication.

References

[1] Han J, Wang S, Yeung Kwong TN, et al. Dermatomyositis as an extrathoracic manifestation of hepatitis B virus-related hepatocellular carcinoma: a case report and literature review. Medicine 2018;97:e11586.
[2] Huber AM, Kim S, Reed AM, et al. Childhood arthritis and rheumatology research alliance consensus clinical treatment plans for juvenile dermatomyositis with persistent skin rash. J Rheumatol 2017;44:110.
[3] Hurliman E, Groth D, Wendelschafer-Crabb G, et al. Small fiber neuropathy in a patient with dermatomyositis and severe scalp pruritus. Br J Dermatol 2017;176:209.
[4] Enomoto Y, Suzuki Y, Hozumi H, et al. Clinical significance of soluble CD163 in polymyositis-related or dermatomyositis-related interstitial lung disease. Arthritis Res Ther 2017;19:9.
[5] Gao MZ, Huang YL, Wu XB, et al. Red blood cell distribution width and neutrophil to lymphocyte ratio are correlated with disease activity of dermatomyositis and polymyositis. J Clin Lab Anal 2018;32:222–209.
[6] Svensson J, Arkema EV, Lundberg IE, et al. Incidence and prevalence of idiopathic inflammatory myopathies in Sweden: a nationwide population-based study. Rheumatology 2017;56:802–10.
[7] Lezhan H, Schade C, Chisolm S, et al. Evidence supports blind screening for internal malignancy in dermatomyositis: data from 2 large US dermatology cohorts. Medicine 2018;97:e9639.
[8] Oztas M, Ugurlu S, Aydin O. Atorvastatin-induced dermatomyositis. Rheumatol Int 2017;37:1217–9.
[9] Baki MA, Donmez S, Saritas F, et al. The epidemiology of dermatomyositis in northwestern Thrace region in Turkey: epidemiology of dermatomyositis in Turkey. Rheumatol Int 2017;37:1–7.
[10] Enders FB, Badermeunier B, Baidam E, et al. Consensus-based recommendations for the management of juvenile dermatomyositis. Ann Rheum Dis 2017;76:329–40.
[11] Satoh M, Tanaka S, Cerveri A, et al. A comprehensive overview on myositis-specific antibodies: new and old biomarkers in idiopathic inflammatory myopathy. Clin Rev Allergy Immunol 2017;52:1–9.
[12] Philbin M, Niewoehner J, Wan GJ. Clinical and economic evaluation of repository corticotropin injection: a narrative literature review of treatment efficacy and healthcare resource utilization for seven key indications. Adv Ther 2017;34:1775–90.
[13] Matsushita T, Mizumaki K, Kano M, et al. Antimelanoma differentiation-associated protein 5 antibody level is a novel tool for monitoring disease activity in rapidly progressive intestinal lung disease with dermatomyositis. Br J Dermatol 2017;176:395.
[14] Huang BY, Chen PC, Chen BH, et al. High-throughput screening of sulfated proteins by using a genome-wide proteome and protein tyrosine sulfation system. Anal Chem 2017;89:3278–84.
[15] Feng Y, Shen Y, Chen H, et al. Expression profile analysis of long noncoding RNA in Acute Myeloid Leukemia by microarray and bioinformatics. Cancer Sci 2017;109:540–53.
[16] Nueda MJ, MartorellMarugan J, Marti C, et al. Identification and visualisation of differential isoform expression in RNA-seq time series. Bioinformatics 2018;34:524–6.
[17] Kustatscher G, Grabowski F, Rappaslbier J. Pervasive coexpression of spatially proximal genes is buffered at the protein level. Mol Syst Biol 2017;13:937.
[18] Belmansour AT, Nourieh R. An aggregation method for performance evaluation of a tandem homogenous production line with machines having multiple failure modes. Reliab Eng Syst Safe 2017;95:1193–201.
[19] Guo J, Li XL, Nie H, et al. Robust luminescent materials with prominent aggregation-induced emission and thermally activated delayed fluorescence for high-performance organic light-emitting diodes. Chem Mater 2017;29:3623–31.
[20] Stankowski S, Sobel JM, Stresfied MA. Geographic cline analysis as a tool for studying genome-wide variation: a case study of pollinator-mediated divergence in a monoflower. Mol Ecol 2017;26:107–22.
[21] Ali RR, Bunting H. Celebrating 25 years of the European society of gene and cell therapy. Human Gene Therapy 2017;28:939.
[22] Xu H, Qiao C, Deng H, et al. Shared relay assignment in cooperative communications for bandwidth maximization. Wireless Networks 2017;23:553–66.
[23] Dadgar S, Wang Z, Johnston H, et al. Asynchronous remodeling is a driver of failed regeneration in Duchenne muscular dystrophy. J Cell Biol 2014;207:139–58.
[24] Bakay M, Wang Z, Melcon G, et al. Nuclear envelope dystrophies show multiple failure modes. Reliab Eng Syst Safe 2017;95:1193–201.
[25] Bakay M, Wang Z, Melcon G, et al. Nuclear envelope dystrophies show multiple failure modes. Reliab Eng Syst Safe 2017;95:1193–201.
[26] Calvo-Pulido MDMP, Martin-Saez E, Fernandez-Guarino M, et al. Asynchronous remodeling is a driver of failed regeneration in Duchenne muscular dystrophy. J Cell Biol 2014;207:139–58.
[27] Calvo-Pulido MDMP, Martin-Saez E, Fernandez-Guarino M, et al. Asynchronous remodeling is a driver of failed regeneration in Duchenne muscular dystrophy. J Cell Biol 2014;207:139–58.
[28] Bakay M, Wang Z, Melcon G, et al. Nuclear envelope dystrophies show multiple failure modes. Reliab Eng Syst Safe 2017;95:1193–201.
Liu et al. Medicine (2020) 99:34

[29] Qi L, Peng X, Huang H, et al. RNA sequencing uncovers the key micro-RNAs potentially contributing to sudden sensorineural hearing loss. Medicine 2017;96:e8837.

[30] Chen X, Duan Q, Xuan Y, et al. Possible pathways used to predict different stages of lung adenocarcinoma. Medicine 2017;96:e6736.

[31] Wang L, You ZH, Xia SX, et al. Advancing the prediction accuracy of protein-protein interactions by utilizing evolutionary information from position-specific scoring matrix and ensemble classifier. J Theor Biol 2017;418:105–10.

[32] Gaspinski M, Luu A, Winter J, et al. CRISP-SKIP; programmable gene splicing with single base editors. Genome Biol 2018;19:107.

[33] Se J, Huang YH, Cui X, et al. Homeobox oncogene activation by pan-cancer DNA hypermethylation. Genomics Biol 2018;19:108.

[34] Ramos M, Schiffer L, Re A, et al. Software for the integration of multihumics experiments in bioconductor. Cancer Res 2017;77:e39.

[35] Varyoska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol 2017;46:1734–9.

[36] Liu YM, You ZH, Xia SX, et al. Advancing the prediction accuracy of protein-protein interactions by utilizing evolutionary information from position-specific scoring matrix and ensemble classifier. J Theor Biol 2017;418:105–10.

[37] Watanabe T. Adult T-cell leukemia: molecular basis for clonal expansion and transformation of HTLV-1-infected T cells. Blood 2017;129:1071.

[38] Guo C, Hou G, Lu X, et al. Mapping protein–protein interactions by double-REDOR-filtered magic angle spinning NMR spectroscopy. J Biomol NMR 2017;67:1–4.

[39] Vortmans T, Ashkenazi A. Cell-cell adhesion: molecular basis for clonal expansion and transformation of HTLV-1-infected T cells. Blood 2017;129:1071.

[40] Pérezsilva JG, Araujovoces M, Quesada V. nVenn: generalized, quasi-proportional Venn and Euler diagrams. Bioinformatics 2018;34:2322–4.

[41] Goncharov T, Hedayan S, Mulvihill MM, et al. Disruption of XIAP–Bcl2 association blocks NOD2-mediated inflammatory signaling. Mol Cell 2018;69:65–61.

[42] Emery IE, Gopalani A, Wood S, et al. Expression and function of ABCG2 and XIAP in ovarian cancer cells. J Neurooncol 2017;133:1–1.

[43] Park CV, Ivanova IG, Kenneth NS. XIAP upregulates expression of HIF target genes by targeting HIF1α for Lys6-linked polyubiquitination. Nucleic Acids Res 2017;45:9336–47.

[44] Li X, Chen W, Zeng W, et al. microRNA-137 promotes apoptosis in ovarian cancer cells via the regulation of XIAP. Br J Cancer 2017;116:66–76.

[45] St HC, Ziegler SG, Markello TC, et al. NT5E mutations and arterial hypertension. Clin Invest 2017;127:1425–37.

[46] Sironi M, Peri AM, Cagliani R, et al. TLR3 mutations in adult patients with Herpes Simplex virus and Varicella Zoster virus encephalitis. J Infect Dis 2017;215:1340.

[47] Kim GT, Cho ML, Park YE, et al. Expression of TLR2, TLR4, and TLR9 in dermatomyositis and polymyositis. Clin Rheumatol 2010;29:273–9.

[48] Ferrera CR, Crow YJ, Gahl WA, et al. DDX58 and classic Singleton-Merten syndrome. J Clin Immunol 2019;39:75–80.

[49] Wahadat MJ, Bodewes I, Maria NI, et al. Type I IFN signature in interferon-positive and interferon-negative patients with primary Sjögren’s syndrome. Ann Rheum Dis 2016;75:721.

[50] Frias MA, James RW, Gerber-Wicht C, et al. Native and reconstituted HDL activate Stat3 in ventricular cardiomyocytes via ERK1/2: role of sphingosine-1-phosphate. Cardiovasc Res 2017;82:313–23.

[51] Maria NI, Steenwijk EC, Ijpma AS, et al. Contrast mapping expression pattern of RNA-sensing receptors TLR7, RIG-I and MDA5 in interferon-positive and interferon-negative patients with primary Sjögren’s syndrome. Ann Rheum Dis 2016;75:721.

[52] Kano F, Matsubara K, Ueda M, et al. Secreted ectodomain of sialic acid-binding Ig-like lectin-9 and monocyte chemoattractant protein-1 (CCL2) in endothelial cells: role in the regulation of cell adhesion and migration. J Periodontol 2017;88:1149–57.

[53] Ashley CL, Glass MS, Abendroth A, et al. Nuclear domain 10 (ND10) and XIAP: a possible role in the regulation of poly(A) microRNA metabolism. J Biol Chem 2018;293:14899–908.

[54] Andrich K, Hegenbart U, Kimmich C, et al. Aggregation of full-length TLR3 or TLR4 activation triggers a cGAS-independent STING pathway to induce host protection. Eur J Immunol 2017;47:2810–21.

[55] Kirches E, Steffen T, Waldt N, et al. The expression of the MSC-marker CD73 and of NF2/Merlin are correlated in meningiomas. J Neurooncol 2017;133:1.

[56] Sher T, Gertz MA. Stem cell transplantation for immunoglobulin light chain amyloidosis. Curr Prob Cancer 2017;41:129–37.

[57] Andrich K, Hegenbart U, Kimmich C, et al. Aggregation of full-length TLR3 or TLR4 activation triggers a cGAS-independent STING pathway to induce host protection. Eur J Immunol 2017;47:2810–21.

[58] Vracko MP, Vernai C, Guscetti F, et al. Homozygous XIAP gene deletion in a child with neonatal diabetes, skeletal anomalies, and immune deficiency. Eur J Pediatr 2017;176:1213–21.

[59] Ishizu A, Tomaru U, Masuda S, et al. Prediction of response to remission induction therapy by gene expression profiling of peripheral blood in Japanese patients with microscopic polyangiitis. Arthritis Res Ther 2017;19:117.

[60] Tan Y, Yang T, Liu P, et al. Association of the OA53 rs1859330/GA genetic polymorphism with severity of enterovirus-71 infection in Chinese Han children. Arch Virol 2017;162:1–9.

[61] Liu Y, Liu P, Liu S, et al. Oligoadenylate synthetase 3 (OAS3) gene polymorphism is associated with severity of EV71 infection in Chinese children. J Clin Virol 2018;101:29–33.

[62] Ashley CL, Glass MS, Abendroth A, et al. Nuclear domain 10 components upregulated via interferon during human cytomegalovirus infection potently regulate viral infection. J Gen Virol 2017;98:1795.

[63] Kano F, Matsubara K, Ueda M, et al. Secreted ectodomain of sialic acid-binding Ig-like lectin-9 and monocyte chemoattractant protein-1 synergistically regenerate transected rat peripheral nerves by altering macrophage polarity. Stem Cells 2017;35:641–53.

[64] Huang T, Jander G. Abscisic acid-regulated protein degradation causes osmotic stress-induced accumulation of branched-chain amino acids in Arabidopsis thaliana. Planta 2017;246:1–10.

[65] Gommersrempel L, Sundstromreml H, Olsson D, et al. An attenuated strain of Mycobacterium avium subsp. duboisii is virulent in BCG-immune mice, indicating that the potential for disease in humans is not a function of virulence in mice. J Virol 2018;82:1150–9.

[66] Sher T, Gertz MA. Stem cell transplantation for immunoglobulin light chain amyloidosis. Curr Prob Cancer 2017;41:129–37.

[67] Andrich K, Hegenbart U, Kimmich C, et al. Aggregation of full-length immunoglobulin light chains from systemic light chain amyloidosis (AL)
patients is remodeled by epigallocatechin-3-gallate. J Biol Chem 2017;292:2328–44.

[79] Long A, Garver KA, Jones SRM. Synergistic osmoregulatory dysfunction during salmon lice (Lepeophtheirus salmonis) and infectious hematopoietic necrosis virus co-infection in sockeye salmon (Oncorhynchus nerka) smolts. J Fish Dis 2019;42:869–82.

[80] Salah Abdoon AS, Giraudelville C, Kandil OM, et al. Maternal recognition of pregnancy and implantation are not associated with an interferon response of the endometrium to the presence of the conceptus in dromedary camel. Theriogenology 2017;90:301.

[81] Forde N, Carter F, Spencer TE, et al. Conceptus-induced changes in the endometrial transcriptome: how soon does the cow know she is pregnant? Biol Reprod 2011;85:144–56.