# Bioinformatics analysis of potential biomarkers associated with diagnosis and treatment of sarcoidosis involved with different tissues

**Mingcheng Wu**
China Medical University
**Yan Zhu**
China Medical University
**Shiyun Han**
China Medical University
**Yan Kou**
China Medical University
**Xiang Li**
China Medical University
**Xing Gao**
China Medical University
**Hongbo Liu** ( hbliu@cmu.edu.cn )
China Medical University
**Rui He**
China Medical University

**Article**

**Keywords:** Bioinformatics, Pulmonary Sarcoidosis, Dermatologic and Lacrimal involvement, Diagnosis and Treatment

**Posted Date:** May 25th, 2022

**DOI:** https://doi.org/10.21203/rs.3.rs-1664432/v1

## Abstract

# Background

Sarcoidosis is a multisystemic granulomatous disease characterized by the formation of non-caseating granulomas, which occurs variably in organs. It is typically prevailing in lungs as well as skin and lacrimal glands. This study was aimed to demonstrate the potential etiology and pathogenesis of sarcoidosis and identify the effective biomarkers associated with the diagnosis and treatment of sarcoidosis.

# Results

A total of 243 DEGs were screened in the pulmonary datasets, 141 and 389 DEGs were acquired through the skin tissue datasets and lacrimal tissue datasets respectively. In lung tissues, DEGs were mainly focused on chemokine-mediated signaling pathway, cellular response to interferon-γ, cell adhesion and chemokine activity. For DEGs in skin and lacrimal tissues were primarily related to interferon-gamma-mediated signaling pathway, immune response, integral component of membrane and MHC class II receptor activity. The result of KEGG enrichment pathway analysis revealed that DEGs were mainly enriched in interferon-gamma-mediated signaling pathway, cell adhesion molecules, Jak-STAT signaling pathway and inflammatory bowel disease. After PPI network analysis, 10 hub genes were identified, respectively.

# Conclusion

This study indicated that utilizing bioinformatic methods to analyze sarcoidosis in variable situations and finding the similarities and differences in molecular mechanism under different conditions are helpful for us to understand pathogenesis of sarcoidosis in a more comprehensive way, laying the foundation for clinical diagnosis and treatment of sarcoidosis.

## Introduction

Sarcoidosis is a rare granulomatous disease involving multi-organ system characterized by the formation of non-caseating granulomas, which occurs all over the world and the incidence rate varies greatly[1]. It can affect all parts of the body, 90% of which are typically prevailing in lungs, followed by skin and eyes. It mainly occurs among young and middle-aged people, and more women than men[2]. The manifestations of sarcoidosis are mainly accompanied with bilateral hilar and mediastinal lymphadenopathy, coughing and dyspnea. Some patients may not have obvious symptoms, and the rate of misdiagnosis is quite high[3]. Dermatologic involvement ranks the second among affected organs. When skin accumulates, two types of skin damage occur: specific and reactive non-specific skin lesions [4]. In addition, 20–50% of cases have ocular involvement with symptoms of granulomatous uveitis and

multifocal choroiditis (MFC) lesions accompanied by venous retinal vasculitis[5]. Due to the variability of initial manifestations and lack of specific and effective screening methods, it can easily lead to deviation in the diagnosis and treatment of sarcoidosis, consequently affecting the prognosis. Despite the widespread use of anti-inflammatory therapies, the morbidity and mortality of sarcoidosis are still on the rise, and glucocorticoid-based therapy may also lead to serious adverse reactions[6]. Some studies have been carried out on sarcoidosis over recent years, but we still know little about the cause of sarcoidosis[7]. It is generally recognized that it might be an autoimmune disease associated with infectious, immune or genetic factors[8]. Therefore, we urgently need to make breakthroughs in the etiology and pathogenesis and find effective biomarkers for the diagnosis and treatment of sarcoidosis.

As an efficient and large-scale genetic data acquisition technology, DNA microarray has been widely applied to the collection of GeneChip expression data and the research of gene expression profiles for many human diseases[9]. These microarray data could provide a new approach to study disease-related genes and also bring great prospect for molecular prediction, drug-based targeting, and treatment[10]. Nowadays, studies related to sarcoidosis have been carried out and published on public databases, which allows deeper researches on molecular mechanisms. For instance, in a research of 37 whole-exome genes in three families, Alain found that, the formation of granulomas may be associated with combination of defects in autophagy and intracellular transport, regulation of G protein, T cell activation, and mitosis and / or immunological synapses[11]. Crouser analyzed the genes expression of pulmonary sarcoidosis to distinguish DEGs between sarcoidosis samples and normal lung samples, and obtained 319 DEGs[12]. However due to various factors, the results of significantly expressed mRNA might be inconsistent. In addition to the involvement in lungs, sarcoidosis also does harm to skin and lacrimal glands, which were rarely studied over the several years. Ocular and skin sarcoidosis are always associated with other organ involvement and can at any stage of the disease. They are also accompanied with psychological impairment and required to achieve best therapy[13, 14].

In order to reveal the etiology and molecular mechanism of sarcoidosis, we analyzed and compared the genes from three different organs. In this study, we discussed the biological functions of the DEGs in different tissues and the similarities and differences of its main signaling pathways, and then confirmed the hub genes that played a major role in the biological process of sarcoidosis in different tissues. This study offered reliable biomarkers for the early screening and helped to reveal the etiology of sarcoidosis as well as its underlying molecular mechanism.

## Materials And Methods

# Selection of gene expression profile

The genes expression data GSE32887, GSE75023, GSE105149 (Table 1) were acquired from GEO database according. GSE32887(platform: GPL570), came from the skin including 26 patients with sarcoidosis and five normal individuals. GPL571, the platform for GSE75023, derived from lung tissues,

containing 15 patients and 12 normal individuals. Microarray data of 8 patients and 7 normal samples, sourced from lacrimal glands, were registered in GSE105149(GPL570).

Table 1
Details for GEO sarcoidosis of three different datasets.

| Involved Parts | GEO | Platform | Control | Case |
|---|---|---|---|---|
| Dermatologic | GSE32887 | GPL570 | 5 | 26 |
| Pulmonary | GSE75023 | GPL571 | 12 | 15 |
| Lacrimal | GSE105149 | GPL570 | 7 | 8 |

# DEGs screening

R was utilized to search DEGs from different sources and we selected robust t-test to screen DEGs between patients with sarcoidosis and normal cases. The adjusted p-value, false discovery rate (FDR), was calculated with by the method of Benjamini□Hochberg. The significantly expressed genes retained for subsequent analysis ought to meet the standard of the absolute log2 fold change |log2FC| > 2 and FDR < 0.05.

# GO and KEGG enrichment analysis

The Database for Annotation, Visualization, and Integrated Discovery (DAVID), an analysis tool for extracting genetic or protein, was used to perform Gene ontology (GO) analysis on the screened DEGs for cellular component (CC), biological process (BP), and molecular function (MF). The Kyoto encyclopedia of genes and genomes (KEGG) database was used to show how genes or other molecules work and relate, and KOBAS is a web server that annotates large gene lists, identifies enriched pathways and associates diseases. By using the KOBAS, KEGG pathway enrichment analysis was performed on DEGs. Only the genes with FDR < 0.05 were considered statistically significant for GO function and KEGG pathway enrichment analysis.

# Construction of PPI network and search for hub genes

The STRING database can identify interactive relationship between known and predicted proteins. When selected DEGs were imported into STRING database, we can obtain the interaction among encoding proteins of DEGs. Each node in the interaction network represents a gene, protein or molecule, and the edges between nodes indicate the interactions of these biomolecules. The proteins closer to the central position are more likely to be core proteins or key candidate genes. The results obtained from the STRING database were imported into Cytoscape software, and the top 10 genes with highest score of the Degree algorithm were considered as hub genes, using the plugin CytoHubba.

## Results

# DEGs screening

Under the cut-off point, | log2FC | > 2 and FDR < 0.05, the results were as follows. In pulmonary parts, a total of 243 DEGs were obtained, including 197 upregulated genes and 46 downregulated genes (because the number of genes selected according to the criteria above was too small, so for this dataset, we use | log2FC | > 1 and FDR < 0.05 as criteria). 141 DEGs in dermatologic and 389 DEGs in lacrimal tissues were screened out. Amid the DEGs, 67 and 288 genes were upregulated genes while 74 and 101 genes were downregulated genes in dermatologic and lacrimal parts, respectively. Figure 1 shows the DEGs for each dataset.

Figure 1 DEGs of three different tissues. (A) Volcano map of GSE75023. (B) Volcano map of GSE32887. (C) Volcano map of GSE105149. The blue spots represented downregulated genes, the red represented upregulated genes.

# GO functional enrichment analysis of DEGs

By using the DAVID online analysis tool, GO functional enrichment analysis was performed on the DEGs and FDR < 0.05 was selected as a statistically significant enrichment. The results of different tissues are illustrated in Fig. 2, Fig. 3 and Fig. 4 respectively. The GO analysis results of the dataset the lungs(GSE75023) showed that in biological process term, upregulated DEGs were mainly involved in cell surface receptor signaling pathways, chemokine-mediated signaling pathways, T cell receptor signaling pathways, regulation of immune responses, and inflammation response, positive regulation of JAK-STAT cascade, interleukin-1 cellular response, cellular response to interferon-γ and other processes; and downregulated DEGs were involved in cell adhesion, cellular response to transforming growth factor-stimulated cell response and other processes; In terms of cellular components, upregulated DEGs were basically enriched in membrane, cytoplasmic membrane, cell surface, and extracellular exosomes, while downregulated DEGs were enriched in plasma membrane, extracellular exosomes, and extracellular space. In terms of molecular function, upregulated DEGs were mainly associated with chemokine activity, receptor activity, CXC chemokine receptor activity, CXCR3 chemokine receptor binding, protein homodimerization activity, CCR chemokine receptor binding and other functions, downregulated DEGs were primarily concentrated on hormone activity and other functions.

Results of the skin(GSE32887) showed that in biological processes, yet upregulated DEGs were mostly focused on immune response, interferon-gamma-mediated signaling pathway, regulation of cell shape, response to cytokine and chemokine-mediated signaling pathway and other processes, but the downregulated DEGs were enriched in ion transport, chloride transmembrane transport, C21- steroid metabolism hormone, bicarbonate transport and tyrosine catabolic and other processes; in cellular component term, upregulated DEGs were mainly enriched in activities, including plasma membrane, neuronal cell body, extracellular region and integral component of membrane, and downregulated DEGs were focused on cellular exosomes, extracellular space, cell surface and apical plasma membrane; in terms of molecular functions, upregulated DEGs were basically enriched in carbohydrate binding and MHC class II receptor activity, while downregulated DEGs were associated with functions such as iron ion binding, fatty acid ligase activity, chloride transmembrane transporter activity, fatty acyl-CoA synthase activity, RNA Polymerase II transcription factor activity and sequence-specific DNA binding.

GO analysis of the lacrimal glands (GSE105149) indicated that in biological process group, upregulated DEGs were mainly involved in pathways, containing immune response, inflammatory response, interferon-γ-mediated signaling pathway, positive regulation of T cell proliferation, and cell adhesion. And for downregulated DEGs, they were mainly involved in the detection of chemical stimuli involved in sensory perception of bitter taste, retinal homeostasis, negative regulation of cysteine-type endopeptidase activity, waterway transport, and defense responses to bacterium; among cellular component, the upregulated DEGs were mainly enriched in the extracellular space, MHC class II protein complex, integral component of plasma membrane, and the external side of the plasma membrane, while the downregulated DEGs were enriched in the extracellular space, extracellular exosomes, basolateral plasma membrane and receptor complex; in term of molecular function, upregulated DEGs were primarily related to MHC class II receptor activity, extracellular matrix structural constituent, heparin binding, receptor activity, and MHC class II proteins complex binding, chemokine activity and other functions, however downregulated DEGs are associated to cysteine-type endopeptidase inhibitor activity, sodium channel regulation activity, aryl esterase activity, urea transmembrane transporter activity, glycoprotein activity combine.

Figure 2 Histogram of GO analysis significant functions of DEGs in GSE75023.

Figure 3 Histogram of GO analysis significant functions of DEGs in GSE32887.

Figure 4 Histogram of GO analysis significant functions of DEGs in GSE105149.

# KEGG pathway analysis of DEGs

KEGG pathway enrichment analysis was conducted on KOBAS online analysis database in order to explore the further genes or molecules relationship of DEGs in three different datasets (the larger the P, the more reliable the enrichment of differential proteins in this pathway is.). The top 20 pathways that significantly enriched were selected for mapping. The results are illustrated in Fig. 5. KEGG pathway analysis indicated that, DEGs in lung parts were mainly enriched in chemokine signaling pathway, cell adhesion molecules (CAMs), Jak-STAT signaling pathway, inflammatory bowel disease (IBD), cytokine-cytokine receptor interaction, hematopoiesis and other pathways. In skin, DEGs were mainly associated with CAMs, IBD, drug metabolism-cytochrome P450, butanoate metabolism, asthma, metabolic pathways and other pathways. DEGs of lacrimal parts were significantly enriched in pathways related to hematopoietic cell lineage, staphylococcus aureus infection, CAMs, rheumatoid arthritis and cytokine-cytokine receptor interaction.

Figure 5 Bubble charts of KEGG pathway analysis of DEGs. (A) GSE75023. (B) GSE32887. (C) GSE105149. Gene number represented number of DEGs in this pathway. Each bubble represented a pathway.

# PPI analysis and hub genes identification

The STRING database was used to establish a PPI network of DEGs in three tissues, in which the nodes represented DEGs and the edges indicated the interaction information between DEGs. Eventually, the PPI

network of the lungs included 242 nodes and 1156 edges and there were 121 nodes and 91 edges in the constructed the skin PPI network (Fig. 6). In the part of lacrimal glands (DEGs criteria, | log2FC | > 2), the PPI network had 378 nodes and 3,538 edges. By importing the previous results of STRIING database into Cytoscape software, the top 10 hub genes with highest scores of each dataset were shown below. In the lungs part, the top 10 hub gens were: IL6, CD28, CCL5, CD2, CXCR3, GZMB, IFNG, CCR5, CXCR4, LCK. The hub genes in the skin were STAT1, ICAM1, IFIT3, GBP1, CD274, IFIH1, IRF9, HLA-E, DDX58, HERC5, HLA-G and in lacrimal glands, the hub genes contained PTPRC, ITGAM, CD86, TYROBP, CCL2, MMP9, FN1, ITGAX, IL1B, TLR8.

Figure 6 The PPI network of DEGs constructed by GSE75023(A), GSE32887(B), GSE105149(C). Each node represented a gene or protein and the edges indicated the interactions of them. The nodes closer to the central position are more likely to be hub genes.

# Discussion

In this study, we found that when sarcoidosis occurs, the upregulated DEGs of pulmonary sarcoidosis were mainly enriched in immune defense-related processes, interferon-γ-mediated signaling pathways, chemokine signaling pathways, positive regulation of the JAK-STAT cascade, chemokine activity, receptor activity, CXC chemokine receptor activity, CXCR3 chemokine receptor binding, protein homodimerization activity and CCR chemokine receptor binding. Similarly, for sarcoidosis involving the skin and lacrimal glands, the upregulated DEGs were also involved in various immune responses, chemokine signaling pathways, chemokine-mediated signaling pathways, interferon-γ-mediated signaling pathways and MHC class II receptor activity. For the DEGs involved in the biological process, there was no great difference among the three tissues. In the term of cellular components, the upregulated and downregulated DEGs of the three types of sarcoidosis were enriched in the membrane and intercellular spaces. And the biggest difference was that pulmonary sarcoidosis was not involved in MHC class II receptor activity.

Studies have shown that MHC, interferon-γ, chemokines were substances that played a key role in the immune process of the body[15]. The results of this study indicated that chemokines and MHC class II can promote immune cells to act at inflammatory sites and perform actively in different involved tissues[16]. Of course, DEGs in the three tissues were all related to the immune system, interferon-γ and vitamin D, which may be used as optional drugs for the treatment of sarcoidosis[17].

KEGG analysis revealed that pathways of the three different involved tissues were mainly related to cytokine-cytokine receptor interactions and chemokine signaling pathways. Cytokines and chemokines are important components of the pathology and physiology of sarcoidosis[18]. Increasing evidence suggested that most of the events leading to inflammation, granuloma formation, and tissue damage could be regulated by these mediators and their receptors[19, 20]. Therefore, the pharmaceutical preparations developed according to the pathogenic mechanism of cytokines may become a powerful approach for the treatment of sarcoidosis.

If not limited to the top 20 meaningful KEGG pathways, it can be also found that sarcoidosis was also associated with NOD-like receptor signaling pathways, cell adhesion molecules, inflammatory bowel disease, leishmaniasis, Jak-STAT signaling pathway, ABC transporter, tuberculosis, Wnt signaling pathway. Researches have shown that abnormal regulation of NOD2 signaling might be related to the pathology of various inflammatory diseases, and its downstream signaling proteins such as inhibitor of apoptosis protein (IAP) have become potential therapeutic targets for sarcoidosis[21]. Changes in cell adhesion molecules are also closely related to the inflammatory response in sarcoidosis[22]. Recent studies suggested that overproduction of inflammatory cytokines, such as interferon gamma (IFN-γ), leading to constituent activation of the JAK-STAT pathway, may be a conserved feature of these diseases. The use of JAK inhibitors can stop these signals and improve symptoms in patients with sarcoidosis[23, 24]. Alveolar macrophage ABCG1 is an important regulator of the occurrence of lung granuloma and inflammation. As a granulomatous disease, pulmonary sarcoidosis is closely related to the ABC transporter pathway[25]. The increased activation of Wnt signaling in sarcoidosis may promote the inflammatory process[26]. In inflammatory bowel disease (IBD), immune diseases such as sarcoidosis are more common and IBD may share some common pathogenic mechanisms with sarcoidosis[27, 28]. Both leishmaniasis and tuberculosis have similar clinical manifestations to sarcoidosis and require to be differentiated. In particular, the relationship between tuberculosis and sarcoidosis has been controversial for a long time, and the studies of the connection between them will also help us to understand the mechanism of sarcoidosis[29, 30].

We also used the STRING database to construct the PPI network of the DEGs in the three datasets. To our surprise, the top 10 hub genes were entirely different, but more or less interrelated. Proinflammatory cytokines are the determinants of inflammatory events leading to sarcoidosis, and IL6 plays an important role in the occurrence of sarcoidosis. And studies have shown that progressive pulmonary sarcoidosis may be a fibroproliferative process triggered by IL-6[31, 32]. The role of CD28 and CD2 in pulmonary sarcoidosis has also been confirmed in studies[33, 34]. Microbial antigens can reduce the expression of tyrosine kinase LCK, which is related to the severity of sarcoidosis, therefore many studies have started to explore how to treat pulmonary sarcoidosis in this regard[35]. The research by Prior C et al showed that elevated levels of IFNg are detectable in the majority of untreated patients with stage II/III pulmonary sarcoidosis, while those with the highest levels appear to have a better chance of being cured completely with corticosteroid therapy[36]. In patients with sarcoidosis, IFNg and the receptor of CXCR3 have been turned out that they could induce elevated levels of Th1 chemokine in biopsy specimens and bronchoalveolar lavage fluid (BALF), demonstrating that IFNg is a biomarker for sarcoidosis[37]. STAT1 plays an important role in innate immunity, and many sarcoidosis genomics studies have proved its functions in the immune response of sarcoidosis, and sarcoidosis may be a disease mediated by STAT1[38, 39]. CCR2 is a chemokine receptor whose polymorphism is associated with its susceptibility and protection of sarcoidosis[40]. In all stages of sarcoidosis, CCL2 and CCL5 are important mediators that could induce migration of CCR1, CCR2, and CCR3-expressing monocytes, as well as CCR5-expressing mast cells[41]. Human leukocyte antigen (HLA-G) may play an critical role in inducing and maintaining immune tolerance, and many studies have also suggested that HLA-G is associated with sarcoidosis[42].

Agostini confirmed that CD86 can modulate T cell responses in sarcoidosis, explaining part of the mechanism of inflammatory responses in sarcoidosis[43]. Stefan's studies have also reported that the formation of macrophages is associated with the upregulated expression of TYROBP, and the activation status of specific macrophage may be a crucial event in the development of sarcoidosis[44]. The increasing activities such as matrix metalloproteinase-9 (MMP9) may influence the initiation and spread of inflammation and then induce the occurrence of sarcoidosis eventually.[45]

In summary, we used bioinformatics to analyze the genome data of sarcoidosis from different tissue samples, and found the differences and similarities between the pathogenesis of sarcoidosis, which provided a new vision for exploring the possible mechanism of sarcoidosis and studying the diagnosis and treatment of sarcoidosis. Of course, our study still had limitations of lack of experimental verification and uncertain results, which could be affected by updating databases. For all this, our study still played a complementary and guiding role in the exploration of sarcoidosis. Subsequently, the results of bioinformatics analysis would be further verified by specific experiments.

# Conclusion

In conclusion, our results identified a series of pathways and DEGs from three different tissues, indicating sarcoidosis was primarily associated with biological processes such as immune response, interferon-γ-mediated signaling, and chemokine signaling. In particular, cutaneous sarcoidosis and ocular sarcoidosis were more closely related to MHC class II receptor activity. Among enriched pathways, the Jak-STAT signaling pathway was of great significance in improving the symptoms of sarcoidosis. Finally, PPI network revealed that despite the hub genes were various, the functions were connected and most of the hub genes were related to the occurrence of sarcoidosis. Using bioinformatics and finding the similarities and differences of sarcoidosis with various involving situations would help us understand the mechanism of sarcoidosis more comprehensively and form guidelines for further diagnosis and treatment of sarcoidosis in the future.

# Abbreviations

MFC multifocal choroiditis

DEGs Differently Expressed Genes

FDR false discovery rate

DAVID Database for Annotation, Visualization, and Integrated Discovery

GO Gene ontology

CC cellular component

BP biological process

| | |
|---|---|
| MF | molecular function |
| KEGG | Kyoto encyclopedia of genes and genomes |
| CAMs | cell adhesion molecules |
| IBD | inflammatory bowel disease |
| IAP | inhibitor of apoptosis protein |
| IFN-γ | interferon gamma |
| BALF | bronchoalveolar lavage fluid |
| HLA-G | Human leukocyte antigen |
| MMP9 | matrix metalloproteinase-9 |
| IL6 | interleukin 6 |
| CD28 | CD 28 molecule |
| CCL5 | C-C motif chemokine ligand 5 |
| CD2 | CD 2 molecule |
| CXCR3 | C-X-C motif chemokine receptor 3 |
| GZMB | granzyme B |
| CCR5 | C-C motif chemokine receptor 5 |
| CXCR4 | C-X-C motif chemokine receptor 4 |
| LCK | LCK proto-oncogene |
| STAT1 | signal transducer and activator of transcription 1 |
| ICAM1 | intercellular adhesion molecule 1 |
| IFIT3 | interferon induced protein with tetratricopeptide repeats 3 |
| GBP1 | guanylate binding protein 1 |
| CD274 | CD 274 molecule |
| IFIH1 | interferon induced with helicase C domain 1 |

| | |
|---|---|
| IRF9 | Interferon regulatory factor 9 |
| HLA-E | major histocompatibility complex, class I, E |
| DDX58 | DExD/H-box helicase 58 |
| HERC5 | HECT and RLD domain containing E3 ubiquitin protein ligase 5 |
| PTPRC | protein tyrosine phosphatase receptor type C |
| ITGAM | integrin subunit alpha M |
| CD86 | CD 86 molecule |
| TYROBP | transmembrane immune signaling adaptor TYROBP |
| CCL2 | C-C motif chemokine ligand 2 |
| FN1 | fibronectin 1 |
| ITGAX | integrin subunit alpha X |
| IL1B | interleukin 1 beta |
| TLR8 | toll like receptor |

# Declarations

### Availability of data and materials

The datasets (GSE32887, GSE75023, GSE105149) supporting the conclusions of this article are available in the Gene Expression Omnibus database, https://www.ncbi.nlm.nih.gov/geo/.

# References

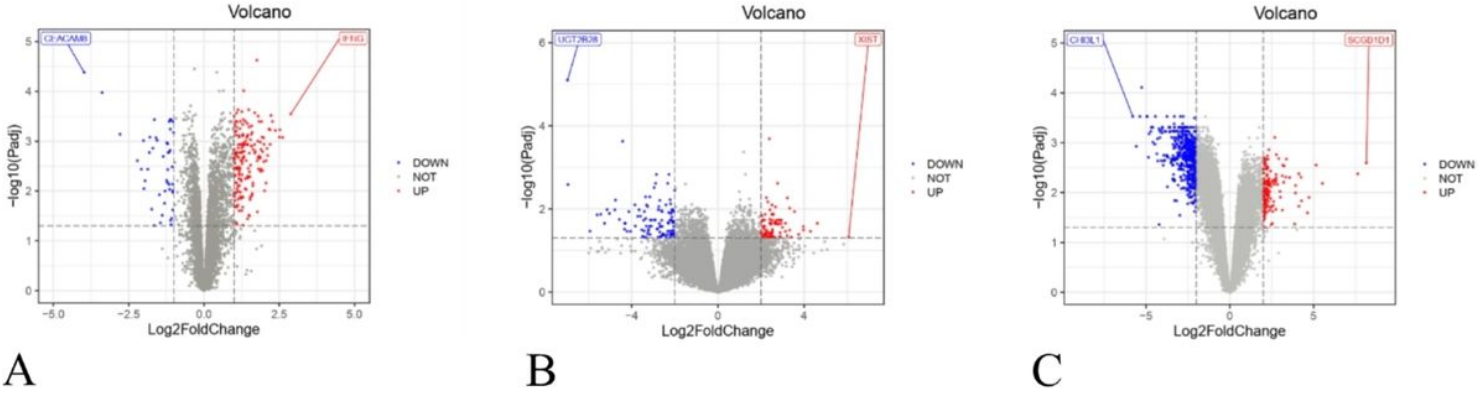

1. Sikjaer MG, Hilberg O, Ibsen R, Lokke A: **Sarcoidosis: A nationwide registry-based study of incidence, prevalence and diagnostic work-up**. Respir Med 2021, **187**:106548.
2. Carmona EM, Kalra S, Ryu JH: **Pulmonary Sarcoidosis: Diagnosis and Treatment**. Mayo Clin Proc 2016, **91**(7):946–954.
3. Bonifazi M, Renzoni EA, Lower EE: **Sarcoidosis and maligancy: the chicken and the egg?** Curr Opin Pulm Med 2021, **27**(5):455–462.
4. Roche FC, Fischer AS, Taylor SC: **Sarcoidosis: An atypical mimicker of acne keloidalis nuchae**. JAAD Case Rep 2020, **6**(5):397–399.
5. Salah S, Abad S, Monnet D, Brezin AP: **Sarcoidosis**. *J Fr Ophtalmol* 2018, **41**(10):e451-e467.

6. Soto-Gomez N, Peters JI, Nambiar AM: **Diagnosis and Management of Sarcoidosis**. Am Fam Physician 2016, **93**(10):840–848.
7. Prasse A: **The Diagnosis, Differential Diagnosis, and Treatment of Sarcoidosis**. Dtsch Arztebl Int 2016, **113**(33–34):565–574.
8. Chen ES, Moller DR: **Etiology of sarcoidosis**. Clin Chest Med 2008, **29**(3):365–377, vii.
9. Petryszak R, Burdett T, Fiorelli B, Fonseca NA, Gonzalez-Porta M, Hastings E, Huber W, Jupp S, Keays M, Kryvych N *et al*: **Expression Atlas update–a database of gene and transcript expression from microarray- and sequencing-based functional genomics experiments**. Nucleic Acids Res 2014, **42**(Database issue):D926-932.
10. Nannini M, Pantaleo MA, Maleddu A, Astolfi A, Formica S, Biasco G: **Gene expression profiling in colorectal cancer using microarray technologies: results and perspectives**. Cancer Treat Rev 2009, **35**(3):201–209.
11. Calender A, Rollat Farnier PA, Buisson A, Pinson S, Bentaher A, Lebecque S, Corvol H, Abou Taam R, Houdouin V, Bardel C *et al*: **Whole exome sequencing in three families segregating a pediatric case of sarcoidosis**. BMC Med Genomics 2018, **11**(1):23.
12. Crouser ED, Culver DA, Knox KS, Julian MW, Shao G, Abraham S, Liyanarachchi S, Macre JE, Wewers MD, Gavrilin MA *et al*: **Gene expression profiling identifies MMP-12 and ADAMDEC1 as potential pathogenic mediators of pulmonary sarcoidosis**. Am J Respir Crit Care Med 2009, **179**(10):929–938.
13. AlRyalat SA, Malkawi L, Abu-Hassan H, Al-Ryalat N: **The impact of skin involvement on the psychological well-being of patients with sarcoidosis**. Sarcoidosis Vasc Diffuse Lung Dis 2019, **36**(1):53–59.
14. Pasadhika S, Rosenbaum JT: **Ocular Sarcoidosis**. Clin Chest Med 2015, **36**(4):669–683.
15. Burke RR, Rybicki BA, Rao DS: **Calcium and vitamin D in sarcoidosis: how to assess and manage**. Semin Respir Crit Care Med 2010, **31**(4):474–484.
16. Mortaz E, Masjedi MR, Tabarsi P, Pourabdollah M, Adcock IM: **Immunopathology of sarcoidosis**. Iran J Allergy Asthma Immunol 2014, **13**(5):300–306.
17. Overbergh L, Stoffels K, Waer M, Verstuyf A, Bouillon R, Mathieu C: **Immune regulation of 25-hydroxyvitamin D-1alpha-hydroxylase in human monocytic THP1 cells: mechanisms of interferon-gamma-mediated induction**. J Clin Endocrinol Metab 2006, **91**(9):3566–3574.
18. Agostini C: **Cytokine and chemokine blockade as immunointervention strategy for the treatment of diffuse lung diseases**. Sarcoidosis Vasc Diffuse Lung Dis 2001, **18**(1):18–22.
19. Li L, Silveira LJ, Hamzeh N, Gillespie M, Mroz PM, Mayer AS, Fingerlin TE, Maier LA: **Beryllium-induced lung disease exhibits expression profiles similar to sarcoidosis**. Eur Respir J 2016, **47**(6):1797–1808.
20. Garman L, Pelikan RC, Rasmussen A, Lareau CA, Savoy KA, Deshmukh US, Bagavant H, Levin AM, Daouk S, Drake WP *et al*: **Single Cell Transcriptomics Implicate Novel Monocyte and T Cell Immune Dysregulation in Sarcoidosis**. Front Immunol 2020, **11**:567342.

21. Tigno-Aranjuez JT, Bai X, Abbott DW: **A discrete ubiquitin-mediated network regulates the strength of NOD2 signaling**. Mol Cell Biol 2013, **33**(1):146–158.
22. Berlin M, Lundahl J, Skold CM, Grunewald J, Eklund A: **The lymphocytic alveolitis in sarcoidosis is associated with increased amounts of soluble and cell-bound adhesion molecules in bronchoalveolar lavage fluid and serum**. J Intern Med 1998, **244**(4):333–340.
23. Damsky W, Thakral D, McGeary MK, Leventhal J, Galan A, King B: **Janus kinase inhibition induces disease remission in cutaneous sarcoidosis and granuloma annulare**. J Am Acad Dermatol 2020, **82**(3):612–621.
24. Wang A, Singh K, Ibrahim W, King B, Damsky W: **The Promise of JAK Inhibitors for Treatment of Sarcoidosis and Other Inflammatory Disorders with Macrophage Activation: A Review of the Literature**. Yale J Biol Med 2020, **93**(1):187–195.
25. McPeek M, Malur A, Tokarz DA, Lertpiriyapong K, Gowdy KM, Murray G, Wingard CJ, Fessler MB, Barna BP, Thomassen MJ: **Alveolar Macrophage ABCG1 Deficiency Promotes Pulmonary Granulomatous Inflammation**. Am J Respir Cell Mol Biol 2019, **61**(3):332–340.
26. Levanen B, Wheelock AM, Eklund A, Grunewald J, Nord M: **Increased pulmonary Wnt (wingless/integrated)-signaling in patients with sarcoidosis**. Respir Med 2011, **105**(2):282–291.
27. Halling ML, Kjeldsen J, Knudsen T, Nielsen J, Hansen LK: **Patients with inflammatory bowel disease have increased risk of autoimmune and inflammatory diseases**. World J Gastroenterol 2017, **23**(33):6137–6146.
28. Fischer A, Nothnagel M, Franke A, Jacobs G, Saadati HR, Gaede KI, Rosenstiel P, Schurmann M, Muller-Quernheim J, Schreiber S *et al*: **Association of inflammatory bowel disease risk loci with sarcoidosis, and its acute and chronic subphenotypes**. Eur Respir J 2011, **37**(3):610–616.
29. Culha G, DoGramaci AC, Hakverd IS, Se CIE, Aslanta SO, Cellk E, Kaya T: **The Investigation of the Association of Cutaneous Leishmaniasis in Biopsy Specimens of the Patients with Granulomatous Disease and Skin Cancer Using the Molecular Method**. Iran J Parasitol 2020, **15**(3):307–314.
30. Gurel MS, Tekin B, Uzun S: **Cutaneous leishmaniasis: A great imitator**. Clin Dermatol 2020, **38**(2):140–151.
31. Bihl MP, Laule-Kilian K, Bubendorf L, Rutherford RM, Baty F, Kehren J, Eryuksel E, Staedtler F, Yang JQ, Goulet S *et al*: **Progressive pulmonary sarcoidosis–a fibroproliferative process potentially triggered by EGR-1 and IL-6**. Sarcoidosis Vasc Diffuse Lung Dis 2006, **23**(1):38–50.
32. Grutters JC, Sato H, Pantelidis P, Ruven HJ, McGrath DS, Wells AU, van den Bosch JM, Welsh KI, du Bois RM: **Analysis of IL6 and IL1A gene polymorphisms in UK and Dutch patients with sarcoidosis**. Sarcoidosis Vasc Diffuse Lung Dis 2003, **20**(1):20–27.
33. Yamaguchi E, Okazaki N, Itoh A, Furuya K, Abe S, Kawakami Y: **Enhanced expression of CD2 antigen on lung T cells**. Am Rev Respir Dis 1991, **143**(4 Pt 1):829–833.
34. Bhargava M, Viken KJ, Barkes B, Griffin TJ, Gillespie M, Jagtap PD, Sajulga R, Peterson EJ, Dincer HE, Li L *et al*: **Novel protein pathways in development and progression of pulmonary sarcoidosis**. Sci Rep 2020, **10**(1):13282.

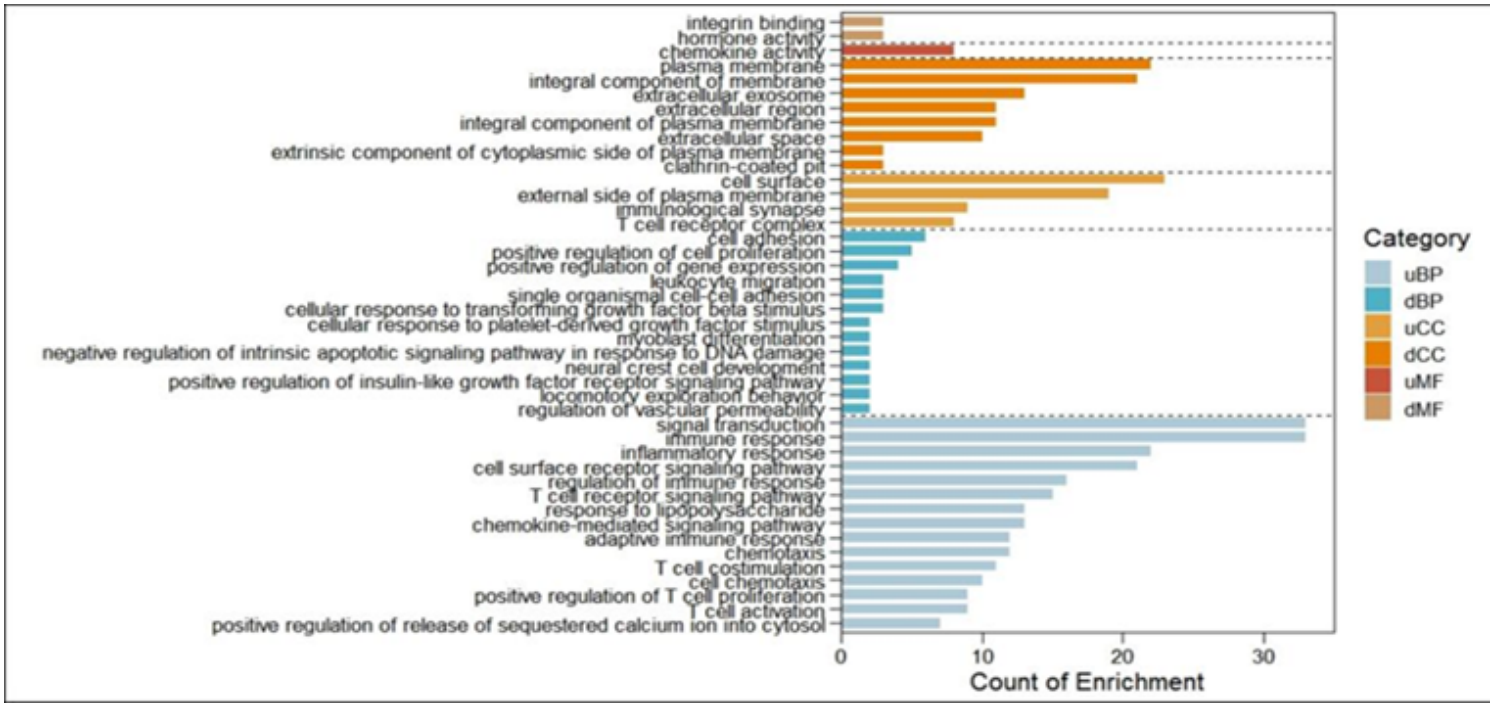
35. Drake WP, Richmond BW, Oswald-Richter K, Yu C, Isom JM, Worrell JA, Shipley GR: **Effects of broad-spectrum antimycobacterial therapy on chronic pulmonary sarcoidosis**. Sarcoidosis Vasc Diffuse Lung Dis 2013, **30**(3):201–211.
36. Prior C, Haslam PL: **Increased levels of serum interferon-gamma in pulmonary sarcoidosis and relationship with response to corticosteroid therapy**. Am Rev Respir Dis 1991, **143**(1):53–60.
37. Giusti C: **Sarcoidosis and the alpha chemokine MIG**. Clin Ter 2020, **171**(2):e161-e166.
38. Rosenbaum JT, Pasadhika S, Crouser ED, Choi D, Harrington CA, Lewis JA, Austin CR, Diebel TN, Vance EE, Braziel RM *et al*: **Hypothesis: sarcoidosis is a STAT1-mediated disease**. Clin Immunol 2009, **132**(2):174–183.
39. Rosenbaum JT, Hessellund A, Phan I, Planck SR, Wilson DJ: **The expression of STAT-1 and phosphorylated STAT-1 in conjunctival granulomas**. Ocul Immunol Inflamm 2010, **18**(4):261–264.
40. Petrek M, Drabek J, Kolek V, Zlamal J, Welsh KI, Bunce M, Weigl E, Du Bois R: **CC chemokine receptor gene polymorphisms in Czech patients with pulmonary sarcoidosis**. Am J Respir Crit Care Med 2000, **162**(3 Pt 1):1000–1003.
41. Palchevskiy V, Hashemi N, Weigt SS, Xue YY, Derhovanessian A, Keane MP, Strieter RM, Fishbein MC, Deng JC, Lynch JP, 3rd *et al*: **Immune response CC chemokines CCL2 and CCL5 are associated with pulmonary sarcoidosis**. Fibrogenesis Tissue Repair 2011, **4**:10.
42. Hviid TV, Milman N, Hylenius S, Jakobsen K, Jensen MS, Larsen LG: **HLA-G polymorphisms and HLA-G expression in sarcoidosis**. Sarcoidosis Vasc Diffuse Lung Dis 2006, **23**(1):30–37.
43. Agostini C, Trentin L, Perin A, Facco M, Siviero M, Piazza F, Basso U, Adami F, Zambello R, Semenzato G: **Regulation of alveolar macrophage-T cell interactions during Th1-type sarcoid inflammatory process**. Am J Physiol 1999, **277**(2):L240-250.
44. Prokop S, Heppner FL, Goebel HH, Stenzel W: **M2 polarized macrophages and giant cells contribute to myofibrosis in neuromuscular sarcoidosis**. Am J Pathol 2011, **178**(3):1279–1286.
45. Piotrowski WJ, Gorski P, Pietras T, Fendler W, Szemraj J: **The selected genetic polymorphisms of metalloproteinases MMP2, 7, 9 and MMP inhibitor TIMP2 in sarcoidosis**. Med Sci Monit 2011, **17**(10):CR598-607.

# Figures



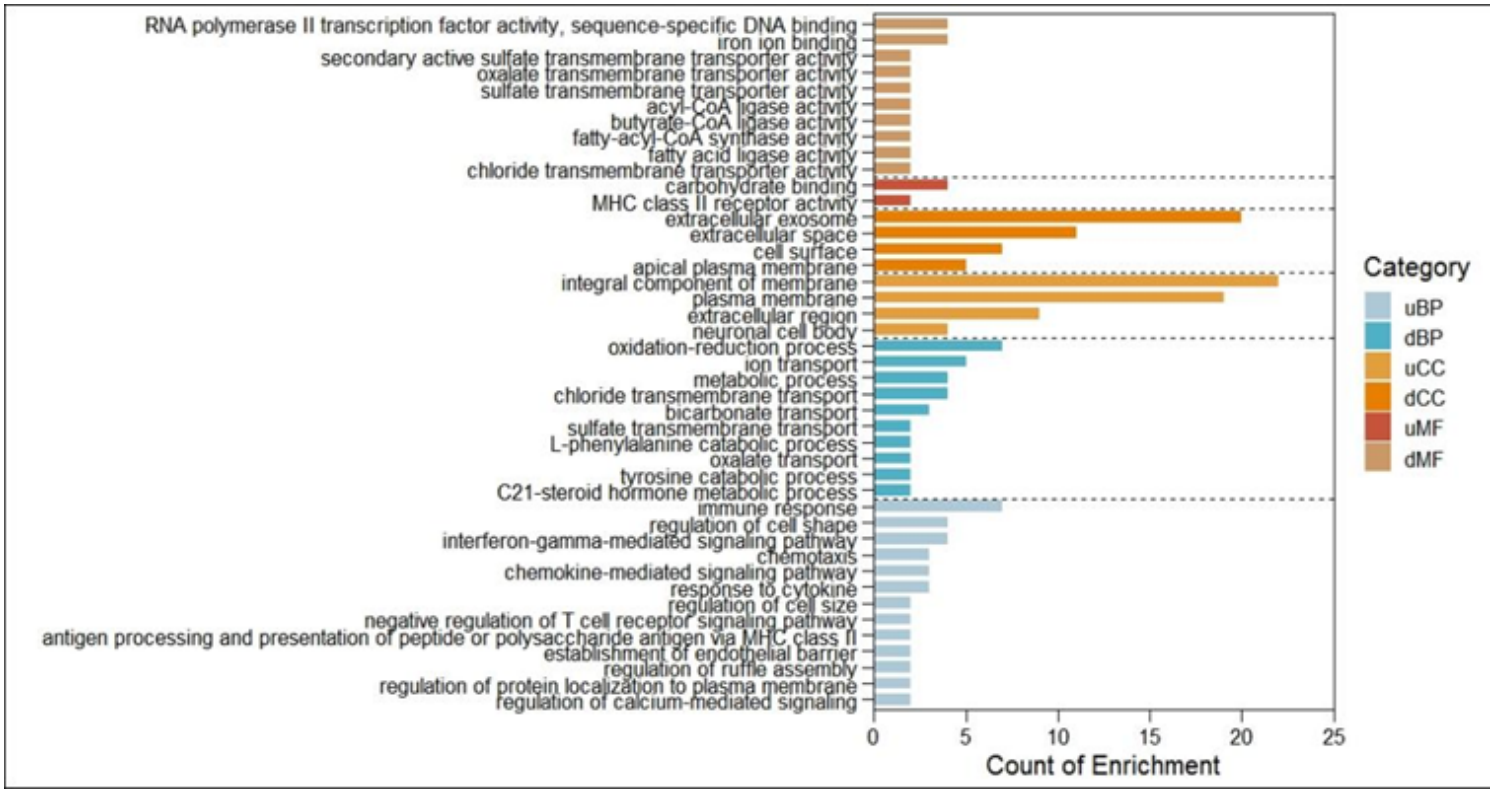
## Figure 1

DEGs of three different tissues. (A) Volcano map of GSE75023. (B) Volcano map of GSE32887. (C) Volcano map of GSE105149. The blue spots represented downregulated genes, the red represented upregulated genes.



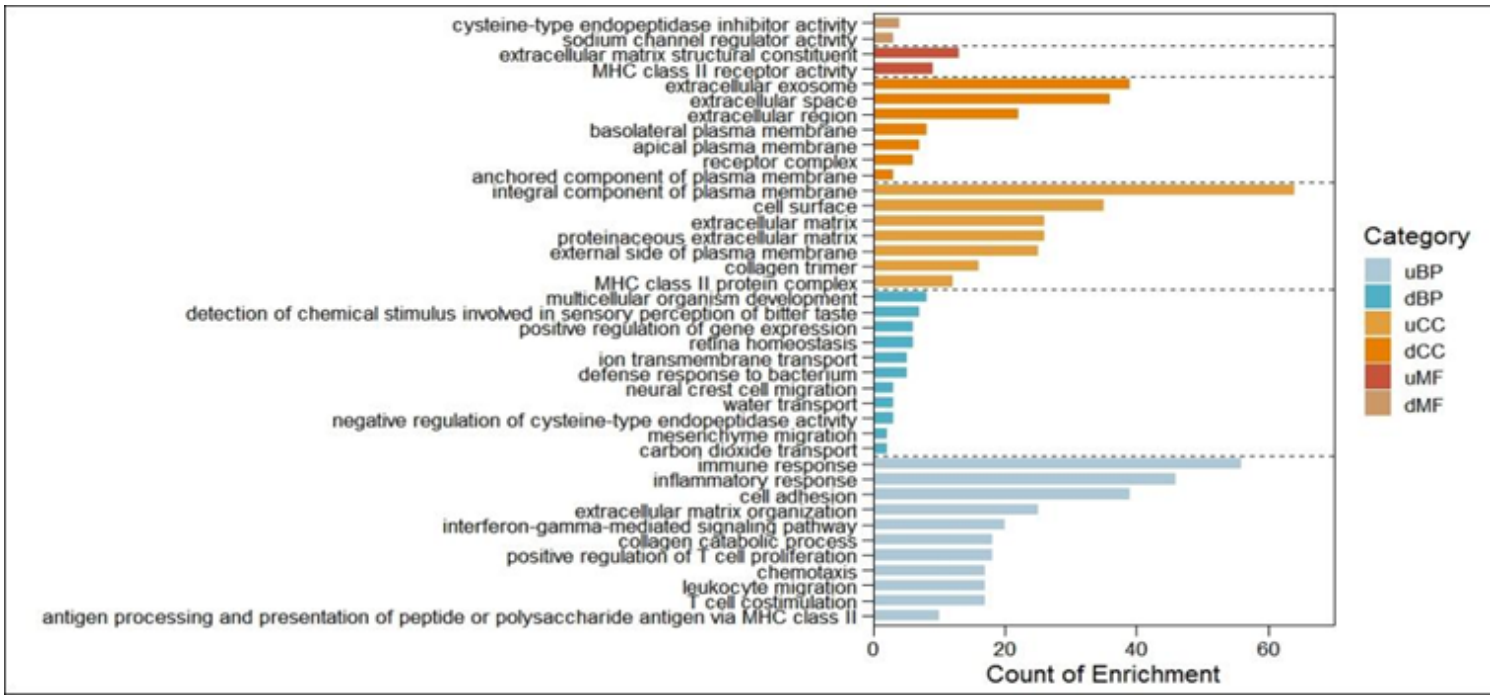

## Figure 2

Histogram of GO analysis significant functions of DEGs in GSE75023.



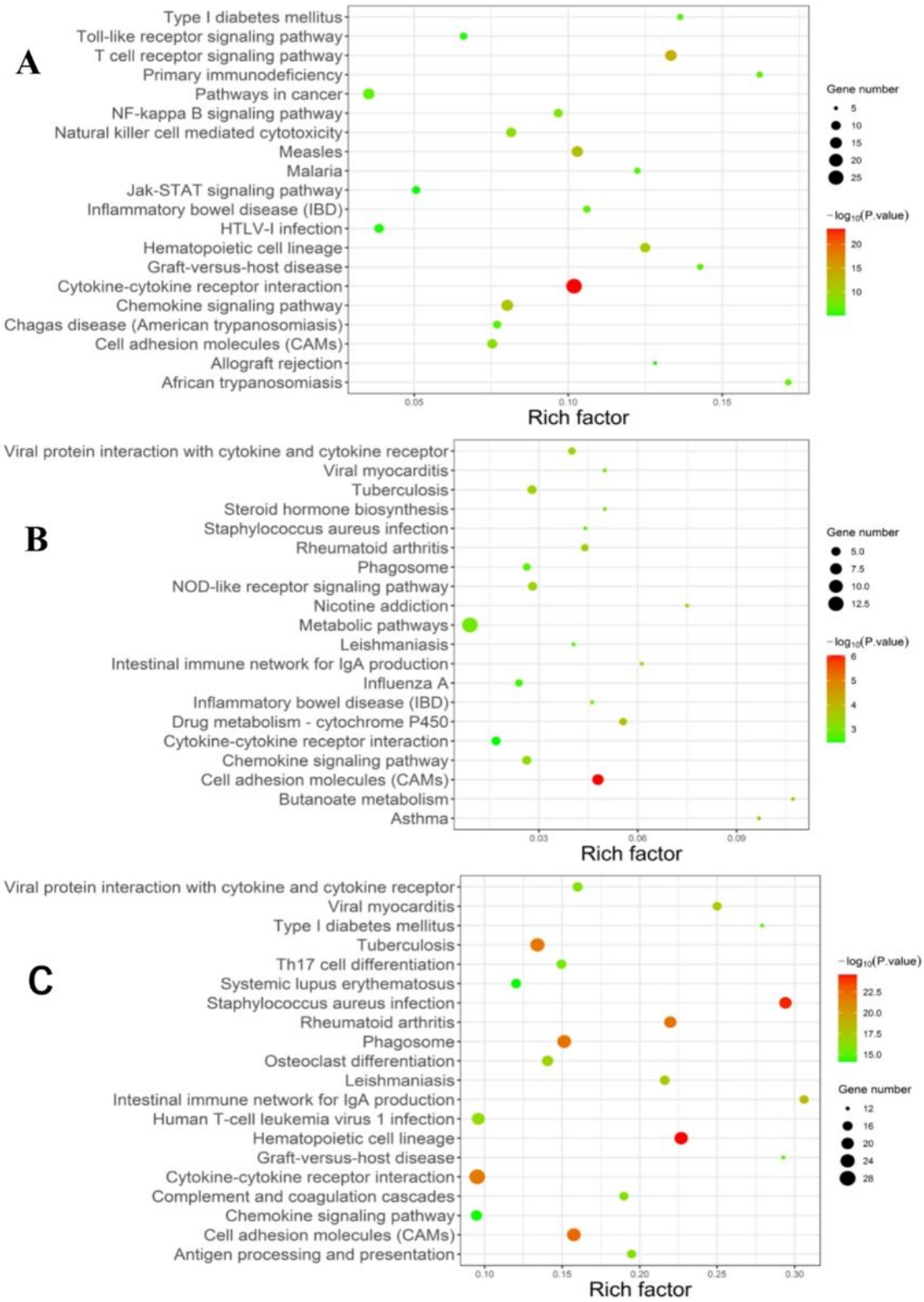

## Figure 3

Histogram of GO analysis significant functions of DEGs in GSE32887.



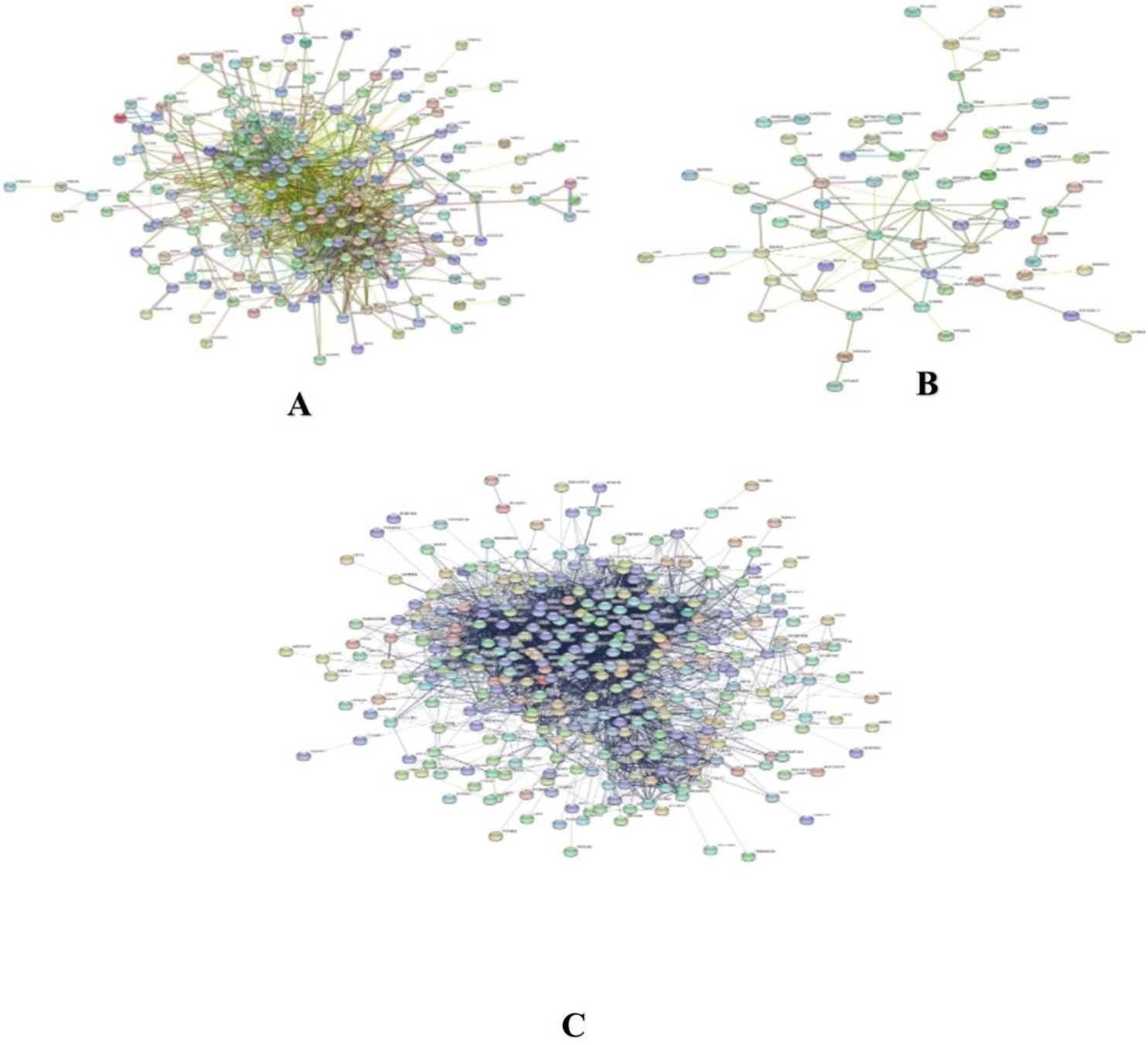

## Figure 4

Histogram of GO analysis significant functions of DEGs in GSE105149.




## Figure 5

Bubble charts of KEGG pathway analysis of DEGs. (A) GSE75023. (B) GSE32887. (C) GSE105149. Gene number represented number of DEGs in this pathway. Each bubble represented a pathway.




## Figure 6

The PPI network of DEGs constructed by GSE75023(A), GSE32887(B), GSE105149(C). Each node represented a gene or protein and the edges indicated the interactions of them. The nodes closer to the central position are more likely to be hub genes.