

Identification of hub-genes, lncRNAs, and skin cancer inducer genes through meta-analysis of Hidradenitis Suppurativa disease transcriptome data.

Mehrdad Shahbazi (✉ me.shahbazi@mail.sbu.ac.ir)

Shahid Beheshti University

Nikta Ziaei

National Institute of Genetic Engineering and Biotechnology

Research Article

Keywords: Hidradenitis Suppurativa, Meta-analysis, LncRNA, Hub-gene, Skin cancer

Posted Date: September 14th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-903999/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

This research is aimed to explore the molecular response of the skin cells to Hidradenitis Suppurativa (HS). Microarray transcriptome data for patients with HS, merged and employed to identify the differentially expressed genes (DEGs), gene ontology (GO), and long non-coding RNAs (lncRNAs). A protein-protein interaction network, and consequently, a co-expression network, was constructed for the essential genes. The key genes' clinical relevance was also established by survival analysis, relative expression level, immunohistochemistry, and immune infiltration correlation analysis. Finally, potential herbal ingredients with inhibitory effects on cancer inducer proteins were characterized using the Traditional Chinese Medicine (TCMD) database. The chemokine signaling pathway was the gene ontology of the most significant cluster found by clusterONE. Myocardial Infarction Associated Transcript (MIAT) and RNA Polymerase II Subunit J4, Pseudogene (POLR2J4) are the predicted lncRNAs for up- and down-regulated genes. Interleukin 6 (IL6), Formyl Peptide Receptor 2 (FPR2), C-X-C Motif Chemokine Ligand 10 (CXCL10), C-C Motif Chemokine Receptor 7 (CCR7), and C-C Motif Chemokine Ligand 5 (CCL5) genes were predicted as hub-genes. Tryptophan 2,3-Dioxygenase (TDO2), Serpin Family B Member 4 (SERPINB4), and Matrix Metalloproteinase 3 (MMP3) were demonstrated to be potential cancer inducers due to their high expression level in HS disease. These genes also were positively correlated with dendritic cells, T cell CD4+, monocytes, and neutrophils in the skin cancer microenvironment. Piceatannol, Salicylic acid, and magnolol herbal ingredients were predicted as potential compounds with inhibitory effects on SERPIN B4, MMP3, and TDO2, respectively. The output of the present study will aid in a better understanding of the HS disease and consequent cancer induction mechanism.

1. Introduction

Hidradenitis suppurative (HS) is a skin disease characterized by painful nodules, abscesses, and draining sinus tracts in the intertriginous area [1, 2]. Patients' quality of life is impaired by this skin condition, and its high frequency results in exorbitant costs [3]. The assumption that microorganisms or skin infections are the first causes of HS has been debunked [4]. The existence of follicular hyperkeratosis has been validated in previous articles to establish the molecular mechanism of HS, and perifolliculitis is considered the earliest occurrence detected in the HS condition [5]. A rupture of the follicular epithelium is the initial inflammatory symptom in HS, followed by a stream of outer body material into the dermis, including bacteria, sebum products, and hair [6]. Bacterial colonization in this environment might induce chronic inflammation [7, 8]. This illness can be treated in various ways, including pharmaceutical and surgical alternatives [2, 9].

Although the molecular link between HS and overall cancer risk is recognized [10–12], the detailed molecular mechanism of cancer induction through this skin condition is a fascinating subject in disease and cancer care. Chronic inflammation and infections, which can lead to proliferative epidermal changes, could have begun the tumor [13, 14]. The higher recurrence of skin malignancies, which are the site of persistent inflammation and bacterial colonization, supports the concept that the illness and neoplasia are connected. Skin cancer has been interpreted to start in chronic wounds, chronic inflammatory

processes, or sclerosing diseases [15, 16]. Genetic background, whole transcriptome analysis, miRNA regulatory elements expression, and cytokine expression analysis have all been employed to elucidate the mechanism of HS at the genomic and proteomic levels [17–19].

Bioinformatics tools could be utilized by researchers to acquire a better knowledge of molecular causes of illness. Microarray data can be leveraged to define a cell's transcript pattern under specific conditions [20]. To characterize the centrality of proteins in a specialized network in response to varied conditions, numerous approaches have been created. Creating different networks, such as PPI and co-expression networks, to lessen the ramifications of biological processes could shed light on the testimonies of the essential genes, product proteins, related pathways, and illness management [21–23]. A considerable amount of transcriptome data is being generated as a result of technological improvements in the life sciences field, which offers a relevant situation for transcriptome data meta-analysis. This technique will provide a more detailed grasp of biological principles and aid in removing false-positive discoveries. The accuracy and reproducibility of the results improved with meta-analysis [24, 25].

In this study, a meta-transcriptome analysis was done on the GSE72702, GSE128637, and GSE148027 datasets, and gene ontology analysis was used to further characterize them. The major clusters, hub-genes, and lncRNAs in response to HS were further identified using a PPI network and centrality analysis. The survival study of the genes in skin cancer and the putative link in skin cancer induction were determined. Finally, potential herbal ingredients with inhibitory effects on the cancer inducer proteins were identified.

2. Methods

2.1. Transcriptome Data collection

For accurate gene expression pattern analysis, three different microarray transcriptome data sets (GSE72702, GSE128637, GSE148027) were downloaded from the NCBI Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) and analyzed to determine the differentially expressed genes among the patients. The GSE72702 included 30 samples, 17 of which were for lesions and 13 of which were healthy. There were a total of 21 samples in the GSE128637, with 10 HS samples and 11 healthy samples. The GSE148027 data comprised 25 samples, 18 of which were related to the lesion and 8 of which were assigned to healthy tissue.

2.2. Preprocessing and (meta) analysis of transcriptome data

Initially, the Linear Models for Microarray Data "LIMMA" package in the R statistical language was used to scan the DEGs in single data analysis [26]. DEGs in each dataset were defined as data with $|\log_2FC$ (fold change) > 2 and adj. P-value < 0.05 . To reflect the common responses across the three datasets, a meta-analysis of transcriptome data was done utilizing the Integrative Meta-Analysis of GEO Data (ImaGEO) web-server following a single data analysis. To integrate the transcriptome data, the maximum P-value

method was utilized, which produces more relevant results than effect size methods. A heat map plot utilizing the ImaGEO web-server displayed the combined data, and a Venn diagram described the up-and down-regulated genes across the three investigations [27].

2.3. GO and pathway enrichment analysis and lncRNA prediction.

The GO of up-and down-regulated genes was represented using the shinyGO server. In this database, the up-regulated genes are also submitted to Reactome and disease-related pathways prediction. The lncRNA prediction was performed for both up-and down-regulated genes [28]. The macro-modules of PPI and built co-expression network genes ontology were supplied via the Metascape service (<https://metascape.org/>). Only data having a significant adj. P-value < 0.05 was considered.

2.4. PPI interaction, module screening, and hub-gene prediction.

A context-based PPI network was built utilizing the STRING database better to define the effective regulation mechanisms among the up-regulated genes. The minimal score for interaction was selected at 0.7 to get more conclusive data from the structured network [29]. To display the built network, cluster characterization, and hub-gene prediction, Cytoscape version 3.7.1 was utilized [30]. Cluster prediction was made via Clustering with the Overlapping Neighborhood Expansion plug-in in Cytoscape [31]. The clusterONE plug-in is a graph clustering method that employs weighted graphs and creates overlapping groups. These features are beneficial for detecting protein complexes in protein-protein interaction networks with related confidence levels. The Hub-gene prediction was conducted using the Cytohubba plug-in with four distinct algorithms, including betweenness, proximity, Maximum Neighborhood Component (MNC), and degree [32].

2.4. Co-expression network of hub-genes and lncRNA

To further describe the role of hub-genes and subsequent pathways, the co-express network was created for hub-genes utilizing the PCViz web server (<http://www.pathwaycommons.org/pcviz>). The co-expression network for predicted lncRNA was generated by COXPRESdb, v 7.3 [33].

2.5. Pathological importance of hub-genes.

Because of the significant correlation between the HS patient and various cancers, pathological analysis was performed. The survival analysis of the hub-genes, top ten up-regulated genes, and lncRNA in skin cancer was described using the Kaplan–Meier plotter database [34]. The expression level of these genes in skin cancer was assessed by the GEPIA database [35]. The immunohistochemistry analysis of the genes in skin cancer was performed through the human protein atlas database [36]. The immune infiltration correlation analysis was achieved through the TIMER web-server [37].

2.6. Drug target prediction

The TCMID database was utilized to characterize the potential herb-based medicines having inhibitory effects on cancer inducer proteins. Using the target search platform in this database, the names of proteins of interest were inserted, and the potential herbal ingredients with inhibitory effects on the proteins' function were collected.

3. Result

3.1. Transcriptome pattern of HS patients.

Microarray data analysis reveals that 2657 DEGs were discovered in GSE128637, 3457 DEGs were discovered in GSE72702, and 7141 DEGs were discovered in the GSE148027 dataset under the FDR 0.05 and adj. P-value < 0.05 thresholds. Figure 1 shows the top three up-and down-regulated genes in each data set. After removing the repeated DEG symbols from the Venn diagram, 34 genes were identified as common DEGs across the three studies. In GSE72702, the most significant expression levels were found in S100 Calcium Binding Protein A7A (S100A7A), SERPINB4, and Immunoglobulin Heavy Locus (IGH) genes. In GSE128637, the Immunoglobulin Lambda Like Polypeptide 1 (IGLL1), S100 Calcium Binding Protein A9 (S100A9), and S100 Calcium Binding Protein A8 (S100A8) genes had the highest expression levels. GSE148027 had the highest expression levels of Fc Receptor-Like 5 (FCRL5), Immunoglobulin Lambda Constant 1 (IGLC1), and Marginal Zone B and B1 Cell Specific Protein (MZB1). An integrative meta-analysis of transcriptome data was performed to further characterize the common DEGs among the three datasets, and the results revealed that SERPINB4, TDO2, ADAM Like Decysin 1 (ADAMDEC1), FCRL5, S100A7A, C-X-C Motif Chemokine Ligand 13 (CXCL13), MMP3, TNF Receptor Superfamily Member 17 (TNFRSF17), S100A9, and C-X-C Motif Chemokine Ligand 6 (CXCL6) genes had the highest expression levels (Fig. 2).

3.2. Gene ontology and pathways

The list of up-and down-regulated genes was analyzed using GO in three categories after merging the data: biological process (BP), molecular function (MF), and cellular component (CC). The disease and Reactome terms associated with the up-regulated genes were investigated further. Some of the main BP GO terms for up-regulated genes include chemokine activity, cytokine activity, Serine type endopeptidase activity, metalloendopeptidase activity, response to biotic stimulus, innate immune response, leukocyte activation, response to cytokine, leukocyte chemotaxis, myeloid leukocyte activation, and response to the bacterium. The immunological synapse, cytoplasmic vesicle membrane, and secretory vesicle are basic terms in CC. In MF, important terms include cytokine binding, immune receptor activity, CCR chemokine receptor binding, CXCR, CXCR3 chemokine receptor binding, and endopeptidase activity. Reactome terms include interleukin signaling, cytokine signaling in immune systems, interleukin ten signaling, GPCR signaling, immune system, and interferon signaling. Macroglobulinemia, synovitis, celiac disease, Crohn's disease, and hyperphosphatemia are some of the most essential terms in disease (Fig. 3). The most downregulated genes in BP were in cell-cell signaling, response to endogenous stimulus, chemical homeostasis, organ development, tissue development, epiderm development, cell differentiation, and skin

development. Cell body, vesicle, cytoskeleton, and myofibrils were some of the CC terms used. Actin, cytoskeletal protein, and steroid-binding were all detected in MF. Ion channel transport, muscle contraction, keratinization, keratinocyte differentiation, epidermal cell differentiation, cornified envelope formation, and cell junction organization were some of the terms found in the Reactome (Fig. 4).

3.3. Cluster, hub-genes, and lncRNA

Context-based PPI was constructed for the up-regulated DEGs as one of the well-known strategies for screening the main clusters and hub-genes. The interactive relationships between the candidate proteins were evaluated using the Cytoscape software. Nine significant clusters (adj. P-value < 0.05) were predicted within the constructed network using the clusterONE plug-in in Cytoscape. Figure 5 depicts the exact genes name. Chemokine-mediated signaling pathway, peptide ligand-binding receptor, calcium ion transport, formyl peptide receptors, G alpha signaling events, cellular calcium ion homeostasis, positive regulation of cytosolic calcium ion concentration, osteoclast differentiation, neutrophil degranulation, osteoclast genesis, Mononuclear cell migration, cellular response to interferon-gamma, regulation of cytosolic calcium ion concentration, was among the most significant GO terms predicted for the clusters. Cytohubba was used to analyze nodes (hub-genes) in the constructed network using four different algorithms: betweenness, closeness, Maximum Neighborhood Component (MNC), and degree. IL6, CCR7, CCL5, CXCL10, and FPR2 are the most critical hub-genes in the network (Table 1). MIAT and POLR2J4 lncRNAs were predicted as the only lncRNAs for up-and down-regulated genes, respectively (Fig. 6).

3.4. Hub-gene and lncRNA co-expression network analysis.

Using the PCViz web server, a co-expression network for the five top hub-genes was created (Fig. 7). Figure 8A shows the GO analysis of the co-expressed genes with hub-genes. Inflammatory response, peptide ligand-binding receptor, interleukin 4, 10, 13, and 17 signaling, response to external stimulus, regulation of leukocyte migration, positive regulation of the MAPK cascade, and positive regulation of cell adhesion are some of the predicted terms. The MIAT lncRNA role in a network for T cell activation, PID IL12 2pathway, regulation of cytokine production, inflammatory bowel disease, cell adhesion molecules, macrophage activation, natural killer cell-mediated cytotoxicity, regulation of DNA binding, regulation of proteolysis, hepatitis C, antigen processing is revealed by the MIAT co-expression network and subsequent GO analysis (Fig. 8).

3.5. Analysis of hub genes, significant DEGs, and lncRNA and from a pathological standpoint.

The survival study of hub-genes was carried out, and as shown in Fig. 10, the increased expression level of hub-genes is directly associated with better survival of skin cancer patients. Among the hub-genes, the expression level of CCL5 and CXCL10 in skin cancer was significantly increased. TDO2, ADAMDEC1, FCRL5, CXCL13, and TNFRSF17 genes, on the other hand, exhibited an up-regulation in skin cancer, whereas the expression level of the MMP3 gene decreased significantly. In skin cancer, the expression

level of MIAT was lower than in normal cells. On the other hand, all the genes showed a better survival rate of patients if they are up-regulated except TDO2, SERPINB4, and MMP3 (Fig. 10). These genes demonstrated a better survival rate of patients if they are down-regulated. Interestingly, in the present study, these genes were up-regulated in the patient's transcriptome data with HS disease. By comparing the expression levels in HS, skin cancer, and survival analysis, three genes, TDO2, SERPINB4, and MMP3, were identified as having potential relevance for further investigation (Table 2). The immunohistochemistry of these three genes in skin cancer is shown in Fig. 11, and the results reveal the up-regulation of the genes in skin cancer. On the other hand, TDO2, SERPINB4, and MMP3 genes were used to identify the correlation between their expression level and the immune infiltration of the skin tumor microenvironment. Dendritic cells, T cell CD4+, T cell CD8+, monocytes, and neutrophils immune cells in TIMER web-server were used, and the positively correlated outputs were considered. These results indicated that the TDO2, SERPINB4, and MMP3 genes might regulate the skin tumor immune microenvironment by affecting the infiltration of immune cells (Fig. 12).

3.6. Predicted herbal ingredients for SERPINB4, MMP3, and TDO2

As a consequence, eleven compounds were predicted for the SERPINB4, MMP3, and TDO2 proteins as they are potential cancer inducers. These substances include Piceatannol, Pinocembrin, and Vitamin E for inhibition of SERPINB4, Magnolol, Sulforathane, and Osthole for inhibition of MMP3, and Salicylic acid (SA), 9-HYROXYETHOXYMETHYLGUANINE (AC2), and Caffeic acid for inhibition of TDO2 (Table 3).

4. Discussion

Transcriptome meta-analysis, identification of hub-genes, lncRNA, and the potential mechanism of skin cancer induction are among the *Insilico* techniques employed in this work. The decrease in false-positive results and the bolding of common responses across studies are two of the most critical advantages of integrating transcriptome data. Linked to the concept of HS illness, the immune system reaction might be due to a later bacterial or viral infection. But the kind of immune response and cancer induction via this disease would enhance the understanding of disease therapy and management and also cancer induction prevention.

In summary, single microarray data functional analysis indicated up-regulation of genes involved in B-cell development/differentiation, inflammatory process control, and immune response. These terms clearly illustrate how an immune response is generated, but they are connected to the illness's second stage, which is a skin infection caused by external stimuli, owing to the HS's disease type. With a significant focus on the up-regulated DEGs as a consequence of combined data, it was demonstrated that genes are engaged in both illness preventive and carcinogenic metabolic pathways. As one of the critical squamous cell carcinoma tumor markers, SERPINB4 is proven to be associated with diverse kinds of skin disorders [38]. TDO2, which is involved in tryptophan amino acid metabolism, has been proven to be a critical component of many malignancies. The exact regulatory mechanism of TDO2 is not precise [39, 40]. The

ADAMDEC1 gene's physiological role is unclear, although it has been proven to be up-regulated in a number of inflammation-related disorders [41, 42]. MMP3 and S100A9 are important genes in the development of cancer and immune responses [43–46]. The FCRL5 gene impacts B-cell development and differentiation, as well as a number of immune-related disorders [47, 48]. As a key gene in the innate immune system, S100A7A plays a function in epidermal differentiation and inflammation [49]. It is widely recognized that this gene has a role in human skin defense against infections. CXCL13 has an influence on inflammatory, infectious, and immunological responses [50, 51]. This gene may also play a role in cancer's start and progression. The protein TNFRSF17 increases B-cell survival and regulates humoral immunity [52]. Both Gram-positive and Gram-negative bacteria have been proven to be sensitive to CXCL6 [53, 54]. Another noteworthy outcome of this study was illness prediction, which was connected to the up-regulated DEGs. Macroglobulinemia is a malignant plasma cell disease in which B cells generate excessive quantities of IgM M-proteins [55]. Lupus is an autoimmune illness in which the body's own tissues, including the skin, are attacked by the immune system [56]. Synovial inflammation occurs when the synovium of a joint becomes inflamed [57]. The immune system assaults the body after consuming gluten, resulting in celiac disease [57]. Crohn's disease is one of the two primary types of inflammatory bowel disease (IBD) [58]. Hyperphosphatemia is an electrolyte condition defined by an abnormally high phosphate concentration in the blood [59]. The majority of people have no symptoms, but some develop calcium deposits in soft tissue. Muscle spasms are frequent as a result of low calcium levels. According to the GO predicted for down-regulated genes, the keratinocyte differentiation metabolism, which is primarily influenced by calcium, and cornification, which is the gradual process of dead cell development on the skin to form a physical barrier, were both dysregulated. In the PPI network that was constructed, the chemokine-mediated signaling pathway obtained the highest score. Chemokine receptor-ligand combinations have a critical function in carcinogenesis, in addition to infection and inflammation [60].

The genes IL-6, FPR2, CCR7, CCL5, and CXCL10 were predicted to behave as hub-genes in response to HS. IL-6 is a pleiotropic cytokine with a broad spectrum of actions in the integrated immune response. One of IL-6's functions is immunological competence, which is defined as a host's capacity to respond to infections [61]. FPR2 identifies formyl peptides generated by pathogens or host cells and contributes to host defense and cell clearance. FPR2 may also identify other ligands with a broad chemical variety that are produced at different phases of inflammation to either promote or resolve inflammation in order to maintain a balanced inflammatory response. The mechanism underpinning FPR2 activation and promiscuous ligand recognition is unknown [62, 63]. CCL5 is categorized as a chemokine or chemotactic cytokine. T cells, eosinophils, and basophils are all chemotactic for CCL5, and it plays a critical role in attracting leukocytes to inflammatory areas [64, 65]. CCR7, a chemokine receptor, is a crucial organizer of the immune system's primary response. CCR7, a CC chemokine receptor, has been identified as a critical regulator of B and T cell trafficking to secondary lymphoid organs in homeostasis [66]. Inflammatory diseases, including viral infections, immunological dysfunction, and tumor development, have been associated with variations in CXCL10 expression levels. CXCL10 is also utilized as a biomarker to predict the severity of a number of illnesses. Infectious illnesses, chronic inflammation, immunological dysfunction, tumor development, metastasis, and dissemination are all connected to CXCL10 in humans

[67, 68]. The predicted up-and-down-regulated DEGs include MIAT and POLR2J4. The MIAT lncRNA is proved to be involved in numerous illnesses and is categorized as a disease-related lncRNA [69]. On the other hand, POLR2J4 as a pseudogene is an RNA polymerase II subunit J4.

The functional enrichment of co-expressed genes with hub-genes and up-regulated lncRNA indicates that both created networks have a tight link to the immune response. GO terms such as cytokine-cytokine receptor interaction, interleukin signaling pathways, T cell activation, and reaction to bacteria were among the most relevant terms. Alongside defining the significant pathways in conjugation with both hub-genes and lncRNA, this built network also provided insight into the characterizing proteins with unknown functions.

In the final part of the study, a combination of gene expression, survival, immunohistochemistry, and immune infiltration studies linked to skin cancer was done. The data indicate that among the hub-genes, CCL5 and CXCL10 genes exhibited a substantial up-regulation in skin cancer. On the other hand, among the up-regulated DEGs of merged data, TDO2, ADAMDEC1, and CXCL13 exhibited up-regulation in skin cancer. But, for correct knowledge of the influence of gene up-or down-regulation on cancer development, survival studies were done for each gene to examine the consistency or inconsistencies linked to the expression level of genes in HS, skin cancer, and survival patterns. Putting together, low expression of TDO2, SERPINB4, and MMP3 proved to enhance the survival rate of patients with skin cancer, and it was in contradiction with examined transcriptome data which indicates the up-regulation of these genes in HS illness. Immunohistochemistry confirmed TDO2, SERPINB4, and MMP3 gene over-expression in skin cancer from various individuals. It has been shown that the immunological state of the tumor microenvironment and the molecular expression of the immune checkpoint can predict the immunotherapy reactivity of patients. It was found that there is a moderate to a significant association between TDO2, SERPINB4, and MMP3 and immune cells. In short, TDO2 exhibited a connection with dendritic cells, T cell CD4+, monocytes, and neutrophils. SERPINB4 indicated a link with T cell CD4+, and MMP3 demonstrated a correlation with monocytes. In malignancies, immune cell infiltration is highly complex, which demands additional in-depth investigation.

Using herbal substances for skincare has been proved to be one of the most reliable strategies due to its favorable molecular effects. To this goal, in the present work, prospective pharmacological targets for SERPIN B4, MMP3, and TDO2 proteins were described using the TCMD database. Through a literature analysis of the anticipated compounds, it was proven that all the compounds have substantial antioxidant activity and, among them, Magnolol, Salicylic acid, and Piceatannol compounds have been demonstrated to be anti-cancer agents [70–72].

5. Conclusion

In conclusion, the present work is intended to undertake a thorough bioinformatics analysis of data relevant to HS illness. IL-6, FPR2, CCR7, CCL5, and CXCL10 were predicted to be hub-genes in response to HS illness. Predicted MIAT lncRNA, together with the stated hub-genes, are expected to be engaged in the

immunological response of the cell against external stimuli. The Top-ten up-regulated DEGs, hub-genes, and predicted lncRNA was subject to clinical significance study, and TDO2, SERPINB4, and MMP3 genes exhibited clinical relevance owing to their up-regulation in HS disease. These genes are also favorably associated with dendritic cells, T cell CD4+, monocytes, and neutrophils. On the other hand, Piceatannol, Salicylic acid, and Magnolol herbal ingredients were demonstrated to be inhibitors of SERPIN B4, MMP3, and TDO2 proteins, respectively. The output of the present work will aid in better defining the HS illness and subsequent cancer induction process.

Declarations

Conflict of Interest

None

References

1. Nguyen, T.V., et al., *Hidradenitis suppurativa: an update on epidemiology, phenotypes, diagnosis, pathogenesis, comorbidities and quality of life*. Journal of the European Academy of Dermatology and Venereology, 2021. **35**(1): p. 50–61.
2. Vinkel, C. and S.F. Thomsen, *Hidradenitis suppurativa: causes, features, and current treatments*. The Journal of clinical and aesthetic dermatology, 2018. **11**(10): p. 17.
3. Krajewski, P.K., et al., *Quality-of-life impairment among patients with hidradenitis suppurativa: a cross-sectional study of 1795 patients*. Life, 2021. **11**(1): p. 34.
4. Revuz, J., *Hidradenitis suppurativa*. Journal of the European Academy of Dermatology and Venereology, 2009. **23**(9): p. 985–998.
5. Alikhan, A., P.J. Lynch, and D.B. Eisen, *Hidradenitis suppurativa: a comprehensive review*. Journal of the American Academy of Dermatology, 2009. **60**(4): p. 539–561.
6. Ralf Paus, L., et al., *What causes hidradenitis suppurativa?* Experimental dermatology, 2008. **17**(5): p. 455–456.
7. Park, Y.J. and H.K. Lee, *The role of skin and orogenital microbiota in protective immunity and chronic immune-mediated inflammatory disease*. Frontiers in immunology, 2018. **8**: p. 1955.
8. Polkowska-Pruszyńska, B., A. Gerkowicz, and D. Krasowska, *The gut microbiome alterations in allergic and inflammatory skin diseases—an update*. Journal of the European Academy of Dermatology and Venereology, 2020. **34**(3): p. 455–464.
9. Zouboulis, C.C., et al., *What causes hidradenitis suppurativa?—15 years after*. Experimental dermatology, 2020. **29**(12): p. 1154–1170.
10. Jung, J.M., et al., *Assessment of overall and specific cancer risks in patients with hidradenitis suppurativa*. JAMA dermatology, 2020. **156**(8): p. 844–853.

11. Makris, G.-M., et al., *Vulvar, perianal and perineal cancer after hidradenitis suppurativa: a systematic review and pooled analysis*. *Dermatologic Surgery*, 2017. **43**(1): p. 107–115.
12. Racanelli, E., et al., *Cutaneous Squamous Cell Carcinoma in Patients with Hidradenitis Suppurativa*. *Cancers*, 2021. **13**(5): p. 1153.
13. Mantovani, A., et al., *Cancer-related inflammation*. *nature*, 2008. **454**(7203): p. 436–444.
14. Perwez Hussain, S. and C.C. Harris, *Inflammation and cancer: an ancient link with novel potentials*. *International journal of cancer*, 2007. **121**(11): p. 2373–2380.
15. MacLeod, A.S. and J.N. Mansbridge, *The innate immune system in acute and chronic wounds*. *Advances in wound care*, 2016. **5**(2): p. 65–78.
16. Badiavas, E.V. and V. Falanga, *Treatment of chronic wounds with bone marrow–derived cells*. *Archives of dermatology*, 2003. **139**(4): p. 510–516.
17. Zouboulis, C.C., et al., *Target molecules for future hidradenitis suppurativa treatment*. *Experimental Dermatology*, 2021. **30**: p. 8–17.
18. Frew, J.W., *Hidradenitis suppurativa is an autoinflammatory keratinization disease: a review of the clinical, histologic, and molecular evidence*. *JAAD International*, 2020. **1**(1): p. 62–72.
19. Dunstan, R.W., et al., *Histologic progression of acne inversa/hidradenitis suppurativa: Implications for future investigations and therapeutic intervention*. *Experimental Dermatology*, 2021. **30**(6): p. 820–830.
20. Dubitzky, W., et al., *Introduction to microarray data analysis, in A practical approach to microarray data analysis*. 2003, Springer. p. 1–46.
21. Liu, G., L. Wong, and H.N. Chua, *Complex discovery from weighted PPI networks*. *Bioinformatics*, 2009. **25**(15): p. 1891–1897.
22. Serin, E.A., et al., *Learning from co-expression networks: possibilities and challenges*. *Frontiers in plant science*, 2016. **7**: p. 444.
23. Xulvi-Brunet, R. and H. Li, *Co-expression networks: graph properties and topological comparisons*. *Bioinformatics*, 2010. **26**(2): p. 205–214.
24. Ahmad, S., et al., *Transcriptome meta-analysis deciphers a dysregulation in immune response-associated gene signatures during sepsis*. *Genes*, 2019. **10**(12): p. 1005.
25. Atay, S., *Integrated transcriptome meta-analysis of pancreatic ductal adenocarcinoma and matched adjacent pancreatic tissues*. *PeerJ*, 2020. **8**: p. e10141.
26. Smyth, G.K., *Limma: linear models for microarray data, in Bioinformatics and computational biology solutions using R and Bioconductor*. 2005, Springer. p. 397–420.
27. Toro-Domínguez, D., et al., *ImaGEO: integrative gene expression meta-analysis from GEO database*. *Bioinformatics*, 2019. **35**(5): p. 880–882.
28. Ge, S.X., D. Jung, and R. Yao, *ShinyGO: a graphical gene-set enrichment tool for animals and plants*. *Bioinformatics*, 2020. **36**(8): p. 2628–2629.

29. Mering, C.v., et al., *STRING: a database of predicted functional associations between proteins*. Nucleic acids research, 2003. **31**(1): p. 258–261.
30. Kohl, M., S. Wiese, and B. Warscheid, *Cytoscape: software for visualization and analysis of biological networks*, in *Data mining in proteomics*. 2011, Springer. p. 291–303.
31. Nepusz, T., H. Yu, and A. Paccanaro, *Detecting overlapping protein complexes in protein-protein interaction networks*. Nature methods, 2012. **9**(5): p. 471–472.
32. Chen, S.-H., et al. *cyto-Hubba: A Cytoscape plug-in for hub object analysis in network biology*. in *20th International Conference on Genome Informatics*. 2009. Citeseer.
33. Obayashi, T., et al., *COXPRESdb v7: a gene co-expression database for 11 animal species supported by 23 co-expression platforms for technical evaluation and evolutionary inference*. Nucleic acids research, 2019. **47**(D1): p. D55-D62.
34. Györfy, B., *Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer*. Computational and Structural Biotechnology Journal, 2021.
35. Tang, Z., et al., *GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses*. Nucleic acids research, 2017. **45**(W1): p. W98-W102.
36. Uhlén, M., et al., *Tissue-based map of the human proteome*. Science, 2015. **347**(6220).
37. Li, T., et al., *TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells*. Cancer research, 2017. **77**(21): p. e108-e110.
38. Izuhara, K., et al., *Squamous cell carcinoma antigen 2 (SCCA2, SERPINB4): an emerging biomarker for skin inflammatory diseases*. International journal of molecular sciences, 2018. **19**(4): p. 1102.
39. Hsu, Y.-L., et al., *Lung cancer-derived galectin-1 contributes to cancer associated fibroblast-mediated cancer progression and immune suppression through TDO2/kynurenine axis*. Oncotarget, 2016. **7**(19): p. 27584.
40. Hattori, S., et al., *Comprehensive behavioral analysis of tryptophan 2, 3-dioxygenase (Tdo2) knockout mice*. Neuropsychopharmacology reports, 2018. **38**(2): p. 52–60.
41. Kumagai, T., S. Fan, and A.M. Smith, *ADAMDEC1 and its role in inflammatory disease and cancer*. Metalloproteinases In Medicine, 2020. **7**: p. 15–28.
42. Liu, T., et al., *ADAMDEC1 promotes skin inflammation in rosacea via modulating the polarization of M1 macrophages*. Biochemical and biophysical research communications, 2020. **521**(1): p. 64–71.
43. Connolly, C., et al., *Enhanced immune response to MMP3 stimulation in microglia expressing mutant huntingtin*. Neuroscience, 2016. **325**: p. 74–88.
44. Biondi, M.L., et al., *MMP1 and MMP3 polymorphisms in promoter regions and cancer*. Clinical Chemistry, 2000. **46**(12): p. 2023–2024.
45. Gebhardt, C., et al., *S100A8 and S100A9 in inflammation and cancer*. Biochemical pharmacology, 2006. **72**(11): p. 1622–1631.
46. Markowitz, J. and W.E. Carson III, *Review of S100A9 biology and its role in cancer*. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2013. **1835**(1): p. 100–109.

47. Davis, R.S., *B-1 cell development and function*. Annals of the New York Academy of Sciences, 2015. **1362**(1): p. 110.
48. Gu, Z., et al., *Genetic risk of FCRL3 and FCRL5 polymorphisms in children with asthma and allergic rhinitis in a Chinese Han population*. International journal of pediatric otorhinolaryngology, 2019. **120**: p. 58–63.
49. Águeda-Pinto, A., L.F.C. Castro, and P.J. Esteves, *The evolution of S100A7: an unusual gene expansion in Myotis bats*. BMC evolutionary biology, 2019. **19**(1): p. 1–11.
50. Kazanietz, M.G., M. Durando, and M. Cooke, *CXCL13 and its receptor CXCR5 in cancer: inflammation, immune response, and beyond*. Frontiers in endocrinology, 2019. **10**: p. 471.
51. Hussain, M., et al., *CXCL13/CXCR5 signaling axis in cancer*. Life sciences, 2019. **227**: p. 175–186.
52. Zhao, C., et al., *POU2AF1, an amplification target at 11q23, promotes growth of multiple myeloma cells by directly regulating expression of a B-cell maturation factor, TNFRSF17*. Oncogene, 2008. **27**(1): p. 63–75.
53. Zhu, Y.M., S. Bagstaff, and P. Woll, *Production and upregulation of granulocyte chemotactic protein-2/CXCL6 by IL-1 β and hypoxia in small cell lung cancer*. British journal of cancer, 2006. **94**(12): p. 1936–1941.
54. Mittal, P., et al., *CXCL6 (granulocyte chemotactic protein-2): a novel chemokine involved in the innate immune response of the amniotic cavity*. American Journal of Reproductive Immunology, 2008. **60**(3): p. 246–257.
55. Treon, S.P., et al., *MYD88 L265P somatic mutation in Waldenström's macroglobulinemia*. New England Journal of Medicine, 2012. **367**(9): p. 826–833.
56. Mok, C. and C. Lau, *Pathogenesis of systemic lupus erythematosus*. Journal of clinical pathology, 2003. **56**(7): p. 481–490.
57. De Lange-Brokaar, B.J., et al., *Synovial inflammation, immune cells and their cytokines in osteoarthritis: a review*. Osteoarthritis and cartilage, 2012. **20**(12): p. 1484–1499.
58. Baumgart, D.C. and W.J. Sandborn, *Crohn's disease*. The Lancet, 2012. **380**(9853): p. 1590–1605.
59. Delmez, J.A. and E. Slatopolsky, *Hyperphosphatemia: its consequences and treatment in patients with chronic renal disease*. American Journal of Kidney Diseases, 1992. **19**(4): p. 303–317.
60. Balkwill, F.R., *The chemokine system and cancer*. The Journal of pathology, 2012. **226**(2): p. 148–157.
61. Chai, Z., et al., *Interleukin (IL)-6 gene expression in the central nervous system is necessary for fever response to lipopolysaccharide or IL-1 beta: a study on IL-6-deficient mice*. The Journal of experimental medicine, 1996. **183**(1): p. 311–316.
62. Hou, X.-L., et al., *FPR2 promotes invasion and metastasis of gastric cancer cells and predicts the prognosis of patients*. Scientific reports, 2017. **7**(1): p. 1–11.
63. Gastardelo, T.S., et al., *Inflammation and cancer: role of annexin A1 and FPR2/ALX in proliferation and metastasis in human laryngeal squamous cell carcinoma*. PloS one, 2014. **9**(12): p. e111317.

64. Aldinucci, D. and A. Colombatti, *The inflammatory chemokine CCL5 and cancer progression*. Mediators of inflammation, 2014. **2014**.
65. Soria, G. and A. Ben-Baruch, *The inflammatory chemokines CCL2 and CCL5 in breast cancer*. Cancer letters, 2008. **267**(2): p. 271–285.
66. Salem, A., et al., *CCR7 as a therapeutic target in Cancer*. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2020: p. 188499.
67. Liu, M., S. Guo, and J.K. Stiles, *The emerging role of CXCL10 in cancer*. Oncology letters, 2011. **2**(4): p. 583–589.
68. Tokunaga, R., et al., *CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation—a target for novel cancer therapy*. Cancer treatment reviews, 2018. **63**: p. 40–47.
69. Alipoor, F.J., M.H. Asadi, and M. Torkzadeh-Mahani, *MIAT lncRNA is overexpressed in breast cancer and its inhibition triggers senescence and G1 arrest in MCF7 cell line*. Journal of cellular biochemistry, 2018. **119**(8): p. 6470–6481.
70. Du, M., Z. Zhang, and T. Gao, *Piceatannol induced apoptosis through up-regulation of microRNA-181a in melanoma cells*. Biological research, 2017. **50**.
71. Ausina, P., et al., *Acetylsalicylic acid and salicylic acid present anti-cancer properties against melanoma by promoting nitric oxide-dependent endoplasmic reticulum stress and apoptosis*. Scientific reports, 2020. **10**(1): p. 1–15.
72. Chilampalli, C., et al., *Effects of magnolol on UVB-induced skin cancer development in mice and its possible mechanism of action*. BMC cancer, 2011. **11**(1): p. 1–17.

Tables 1-3

Tables 1-3 are available in the Supplementary Files section.

Figures

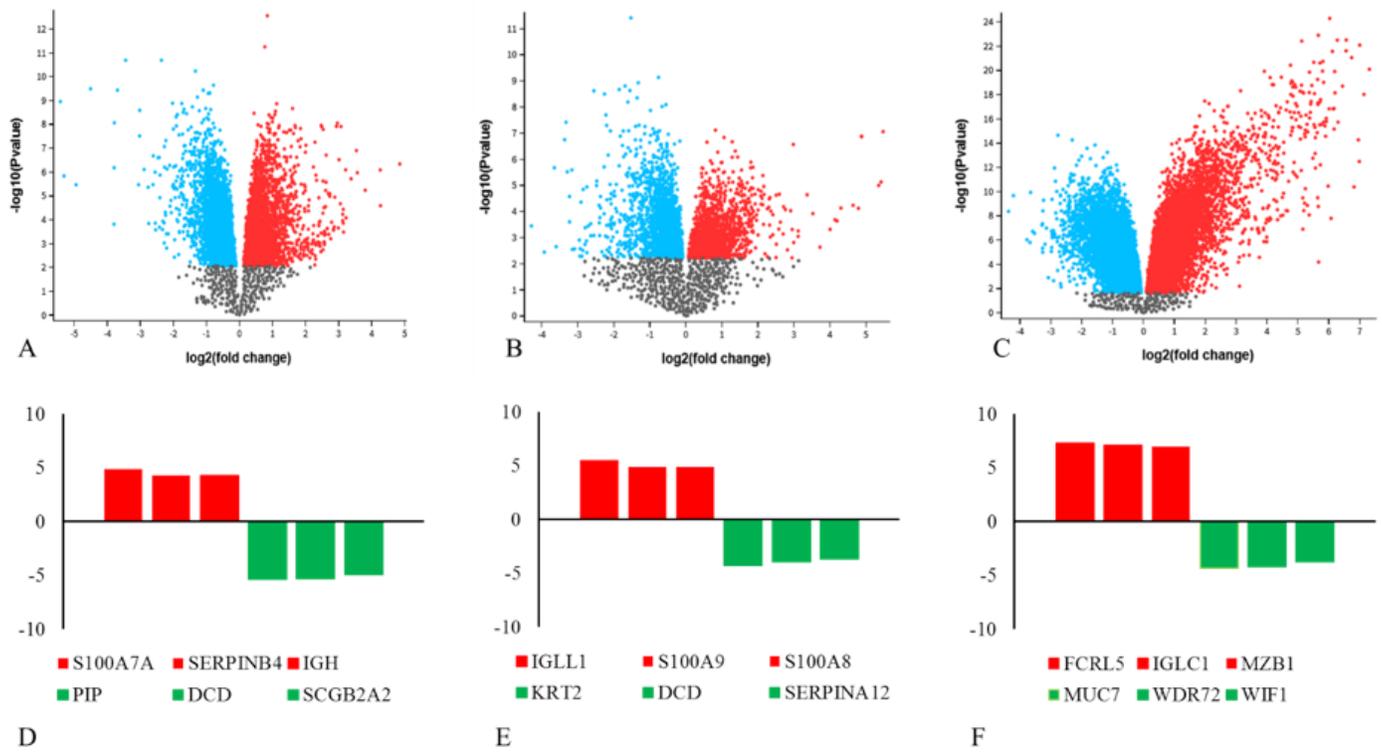
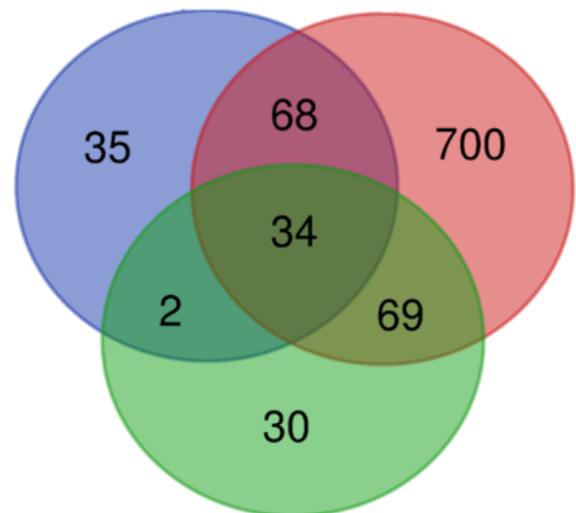
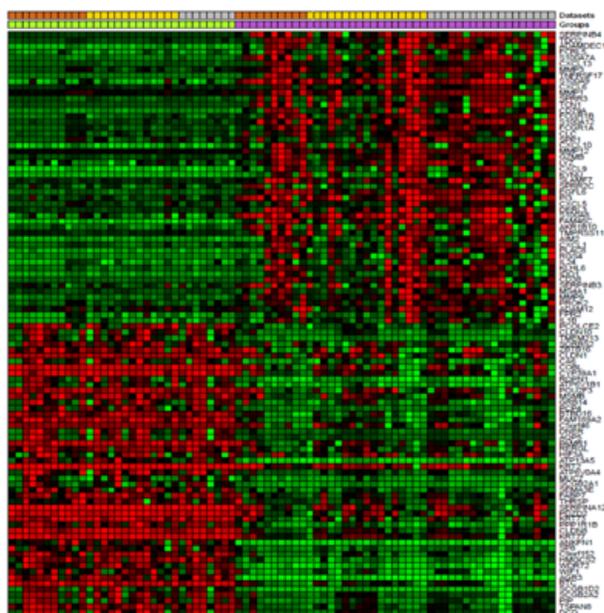


Figure 1

Volcano plots showing the DEGs in three different studies, A) GSE72702 B) GSE128637 C) GSE148027. Red and blue colors demonstrate up and down-regulated genes, respectively. The x-axis demonstrates the fold change, and the y axis demonstrates the $-\log_{10}$ P-values. The expression level of three top up-and down-regulated genes in each study. D) GSE72702, E) GSE128637, F) GSE148027, the x-axis demonstrates the fold change. The adjusted P-value was considered < 0.05 .



A

B

Figure 2

A) Heatmap plot demonstrating the expression level of DEGs in merged data, red and green colors demonstrate the up and down-regulated (log2FC) genes, respectively. The x-axis demonstrates the data sets and samples; the y axis demonstrates the gene symbols. B) Venn diagram (<https://bioinfogp.cnb.csic.es/tools/venny/>) of overlapping DEGs. At least 173 DEGs were overlapped in three data sets.

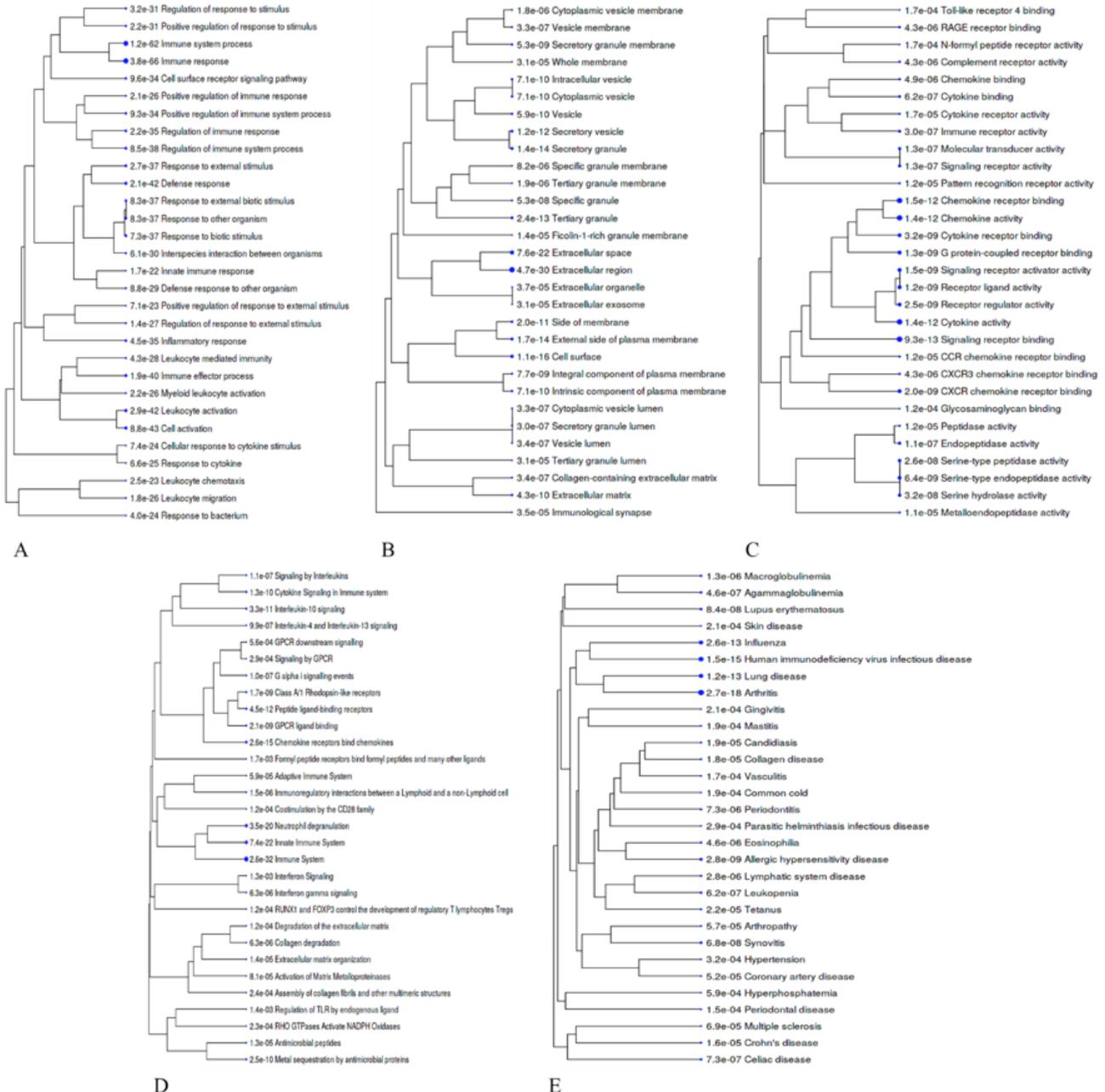


Figure 3

Gene ontology enrichment analysis of up-regulated DEGs of merged data using shinyGO database(<http://bioinformatics.sdstate.edu/go/>). A Molecular function, B, Cellular component, C, Biological process, D, Reactome, E, Disease.

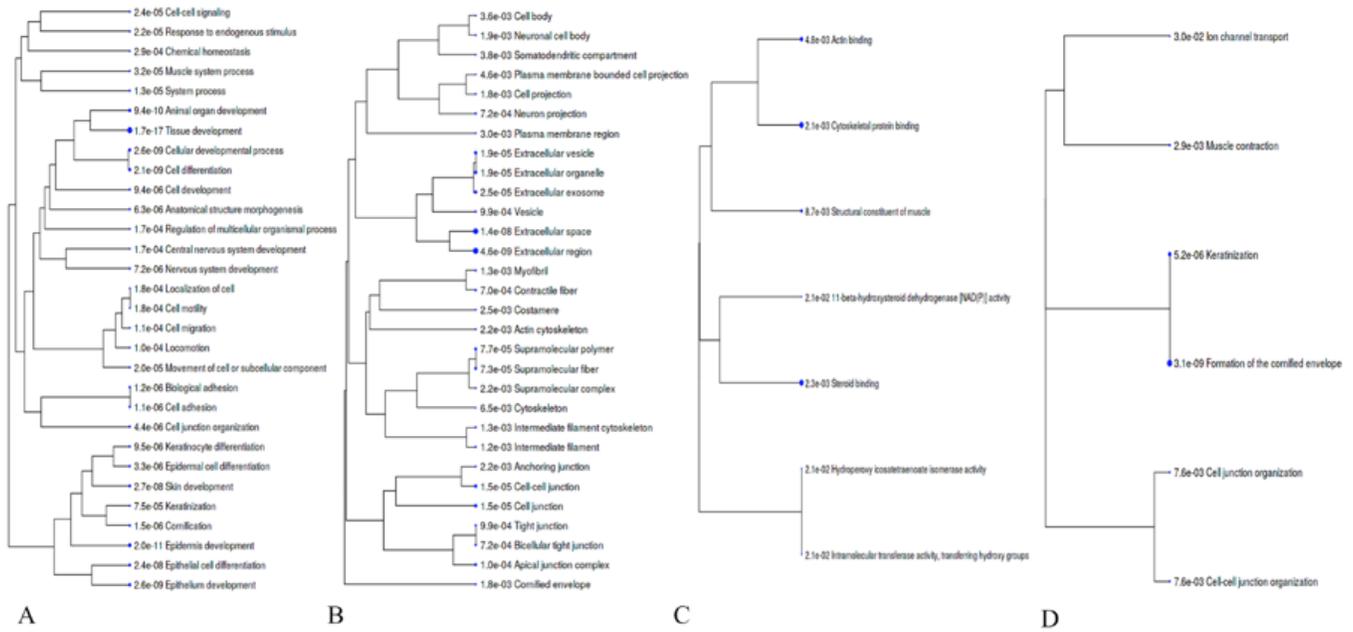


Figure 4

Gene ontology enrichment analysis of down-regulated DEGs of merged data using shinyGO database (<http://bioinformatics.sdstate.edu/go/>). A Molecular function, B, Cellular component, C, Biological process, D, Reactome.

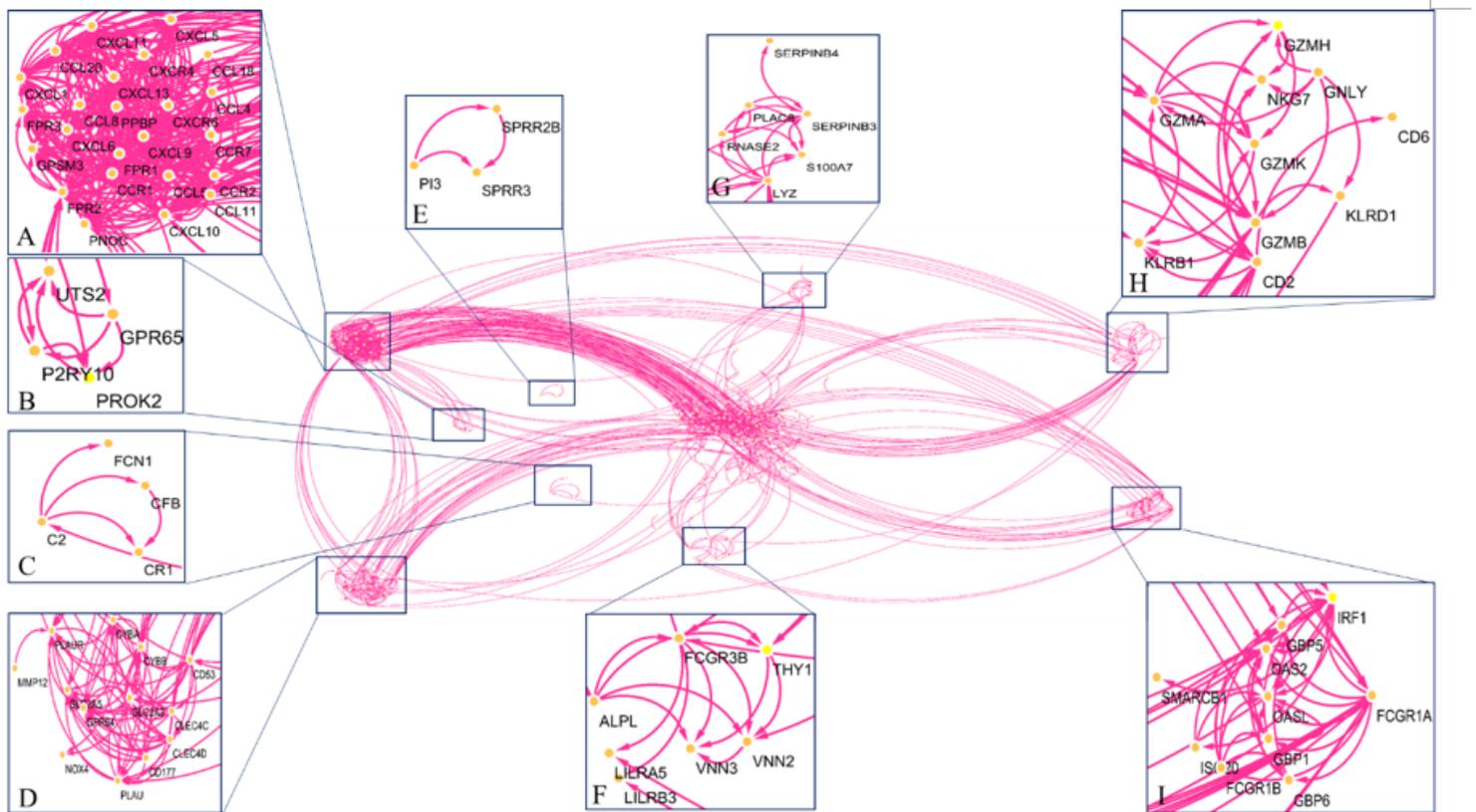


Figure 5

Macro-modules prediction of constructed PPI network. Using the clusterONE plugin, nine significant modules were predicted in the constructed network.

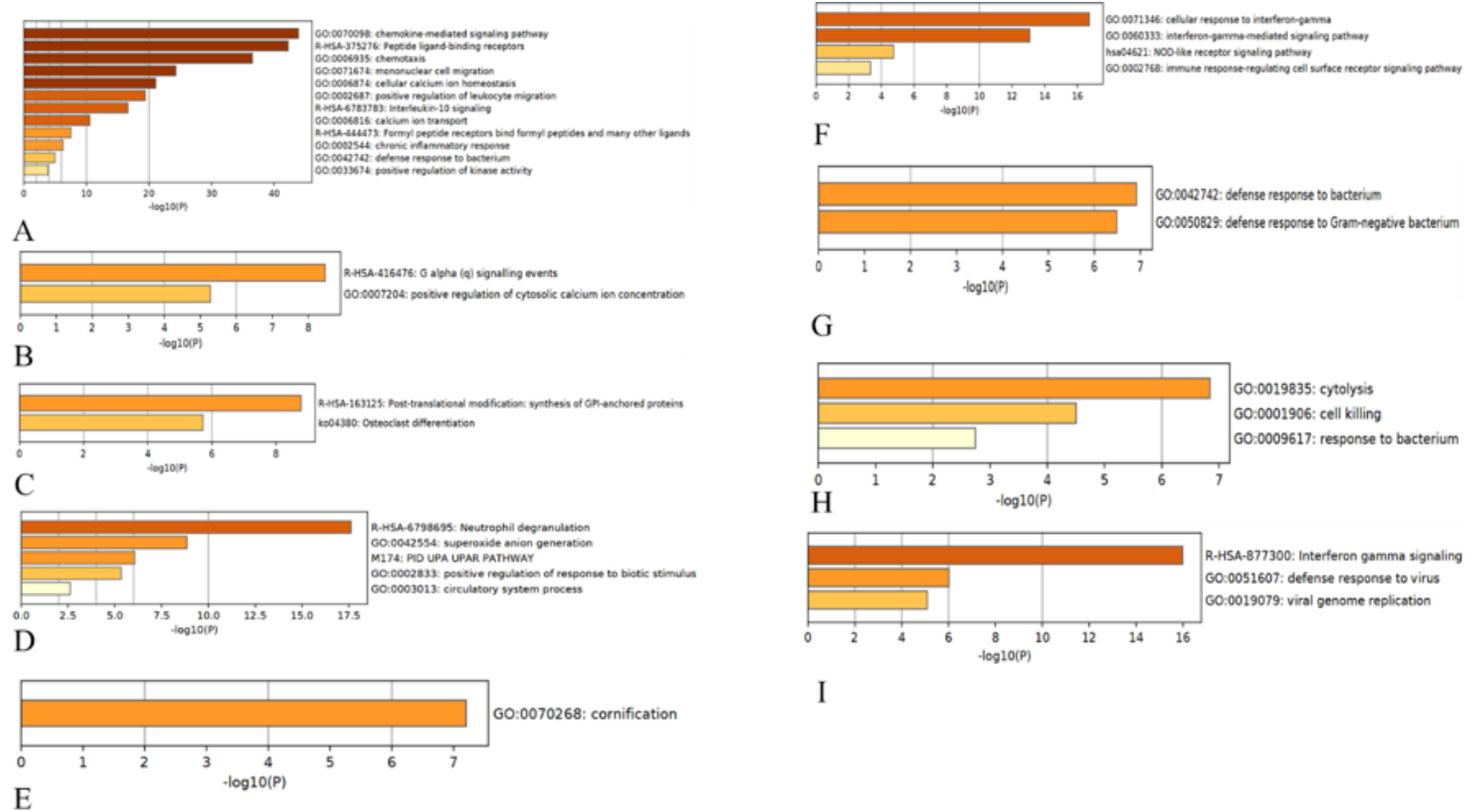


Figure 6

GO Enrichment analysis of predicted macro-modules using (<https://metascape.org/>). The x-axis demonstrates the $-\log_{10}(P\text{-value})$. Bar graph of enriched terms among the hub and co-expressed genes colored based on the P-values.

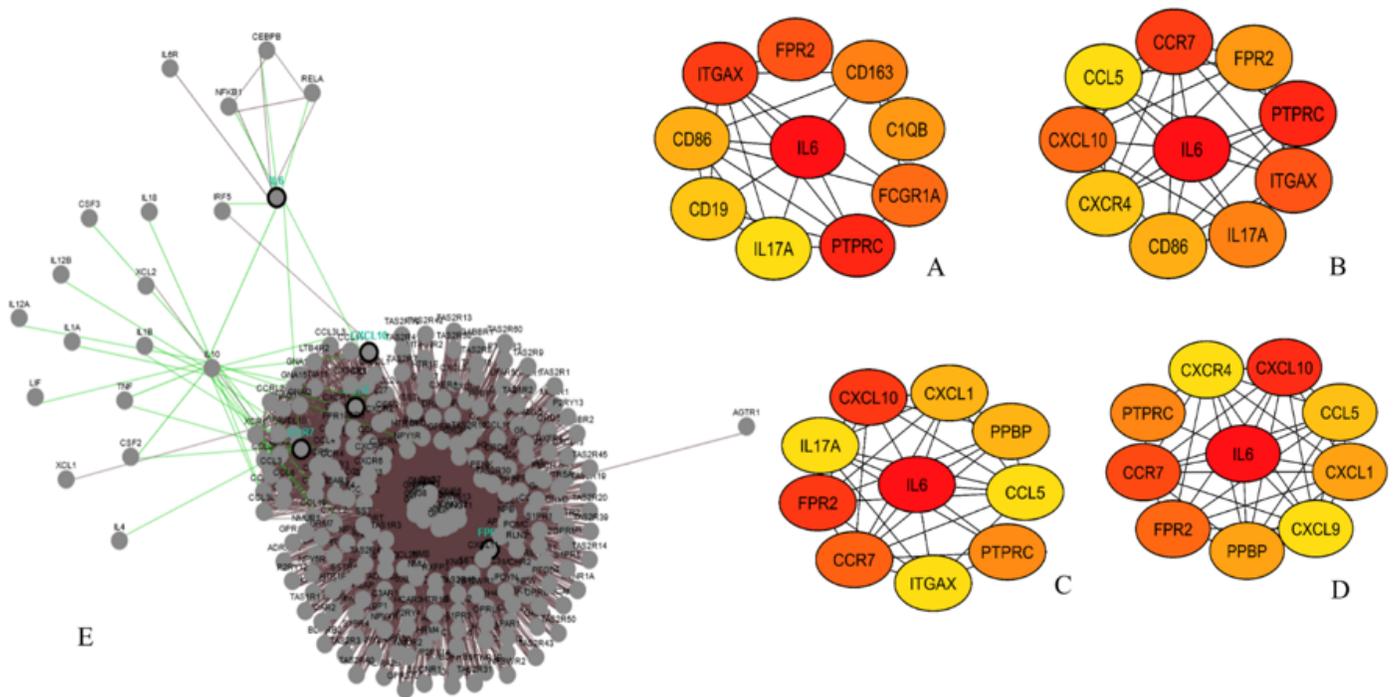


Figure 7

Top ten predicted hub-genes with four different algorithms using cytohubba Cytoscape plugin, A) MNC B) Degree C) Betweenness D) Closeness, E) constructed co-expression network (<http://www.pathwaycommons.org/pcviz>) for the most significant hub genes include, IL6, CXCL10, FPR2, CCR7, CCL5. Genes with the highest number of connections in the network are detected as hub genes. The image shows the degree of importance of hubs through a color scale ranging from red (the most critical) to yellow (less critical).

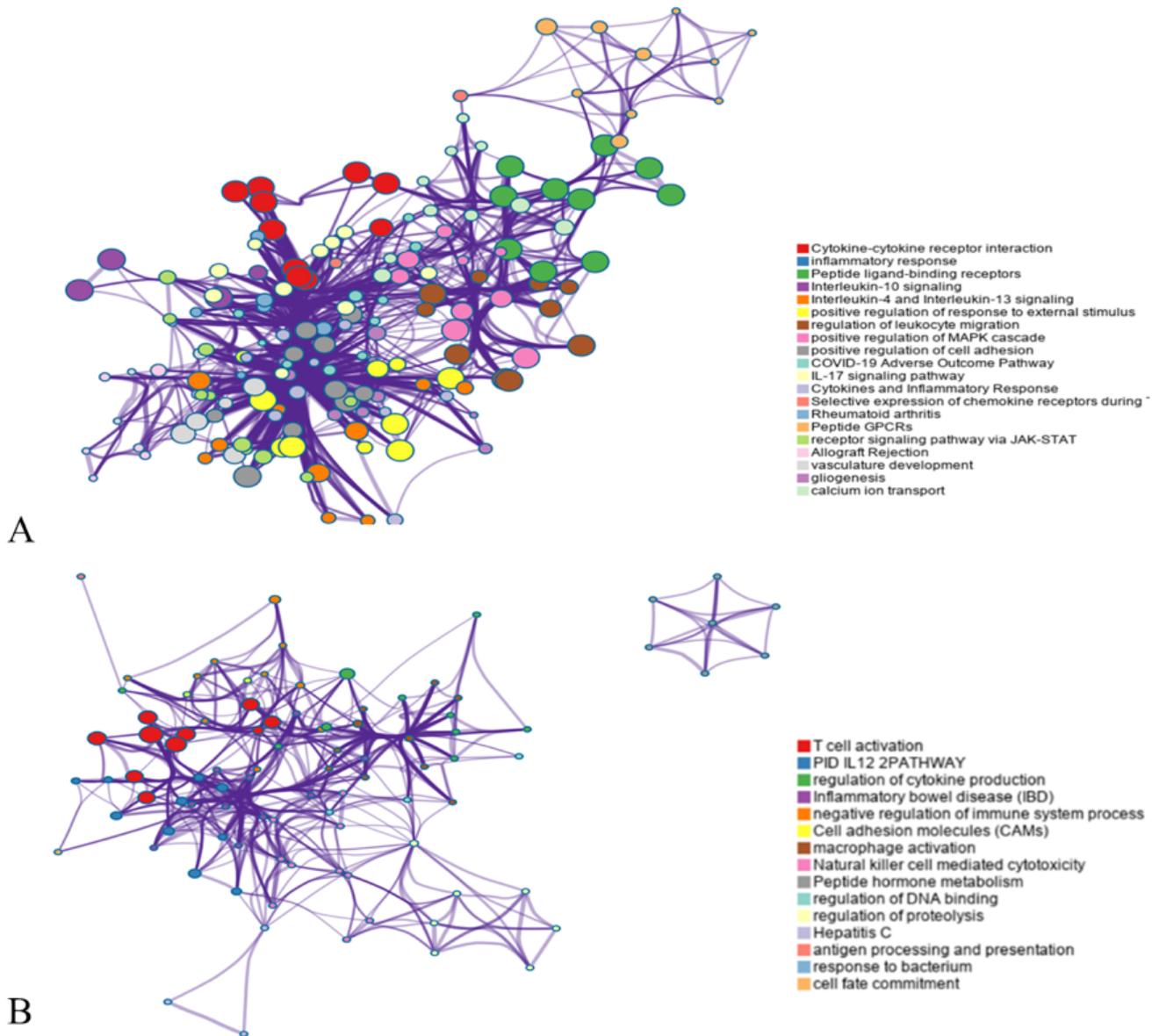


Figure 8

Gene ontology enrichment analysis of constructed co-expression networks (<https://metascape.org/>). A) Constructed co-expression network for top-5 hub-genes. B) GO analysis of constructed co-expression network for predicted up-regulated lncRNA (MIAT).

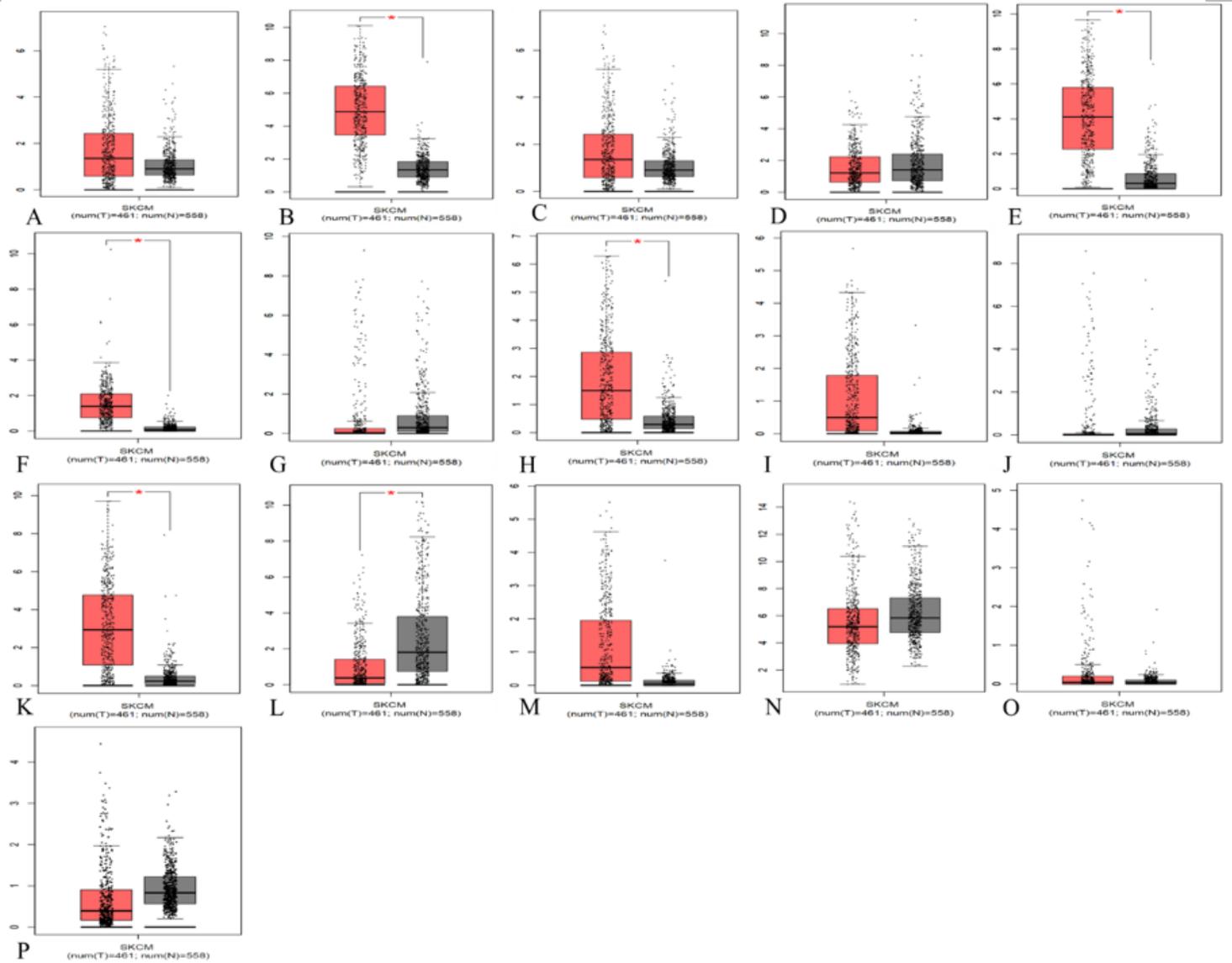


Figure 9

The expression level of hub-genes in skin cancer, using TCGA data in GEPIA. A) FPR2, B) CCL5 C)CCR7 D)IL6 E)CXCL10, F) TDO2, G)SERPINB4, H) ADAMDEC1, I) FCRL5, J) S100A7A, K) CXCL13, L) MMP3, M) TNFRSF17, N) S100A9, O) CXCL6, P) MIAT.

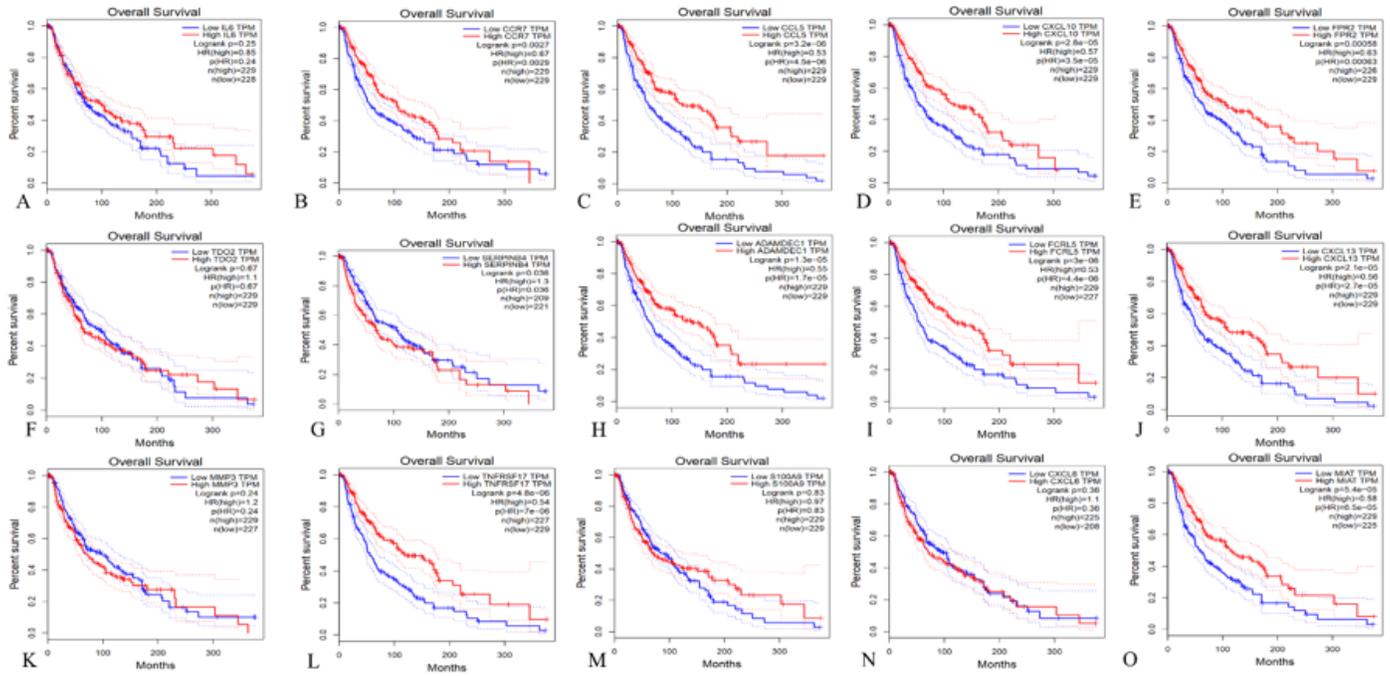


Figure 10

Survival analysis of hub-genes in skin cancer. A) IL6, B) CCR7 C) CCL5 D) CXCL10, E) FPR2 F) TDO2, G)SERPINB4, H) ADAMDEC1, I) FCRL5, J) CXCL13, K) MMP3, L) TNFRSF17, M) S100A9, N) CXCL6, O) MIAT, The red line represents the samples with high gene expression, and the blue line indicates the samples with low gene expression.

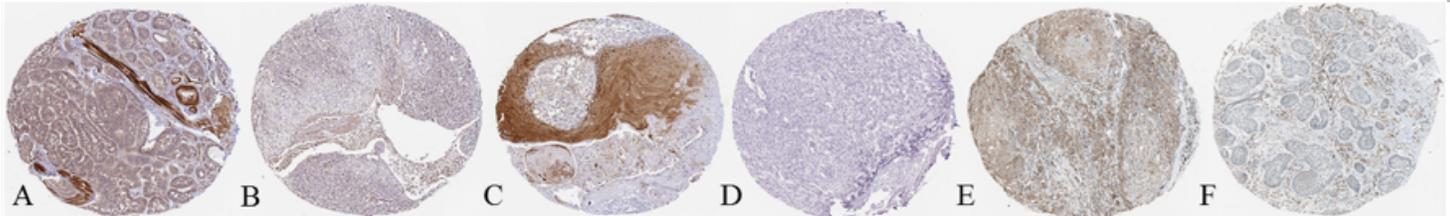


Figure 11

Immunohistochemistry and expression level analysis of TDO2, SERPINB4, and MMP3 genes in skin cancer. A) HPA039611, Male, age 62, Skin (T-01000), Basal cell carcinoma (M-80903), Patient id: 3407, Tumor cells Staining: Medium Intensity: Moderate Quantity: >75% Location: Cytoplasmic/membranous, B) control, HPA007875 Male, age 87 Skin (T-01000) Papilloma, NOS (M-80500) Basal cell carcinoma (M-80903) Patient id: 2548. C) Skin cancer, HPA055992, Female, age 44, Vulva (T-80100), Skin (T-01000), Squamous cell carcinoma, NOS (M-80703), Patient id: 942, Tumor cells, Staining: High Intensity: Strong Quantity: 75%-25% Location: Cytoplasmic/Membranous nuclear, D) HPA048341, Female, age 78, Skin (T-01000), Basal cell carcinoma (M-80903), Patient id: 852. E) Skin cancer, HPA007875, Male, age 75, Skin (T-01000), Lymph node (T-08000), Squamous cell carcinoma, metastatic, NOS (M-80706), Squamous cell carcinoma, NOS (M-80703), Patient id: 2111, Tumor cells, Staining: High, Intensity: Strong, Quantity:

>75%, Location: Cytoplasmic/ membranous, F) HPA007875, Male, age 87, Skin (T-01000), Papilloma, NOS (M-80500), Basal cell carcinoma (M-80903), Patient id: 2548

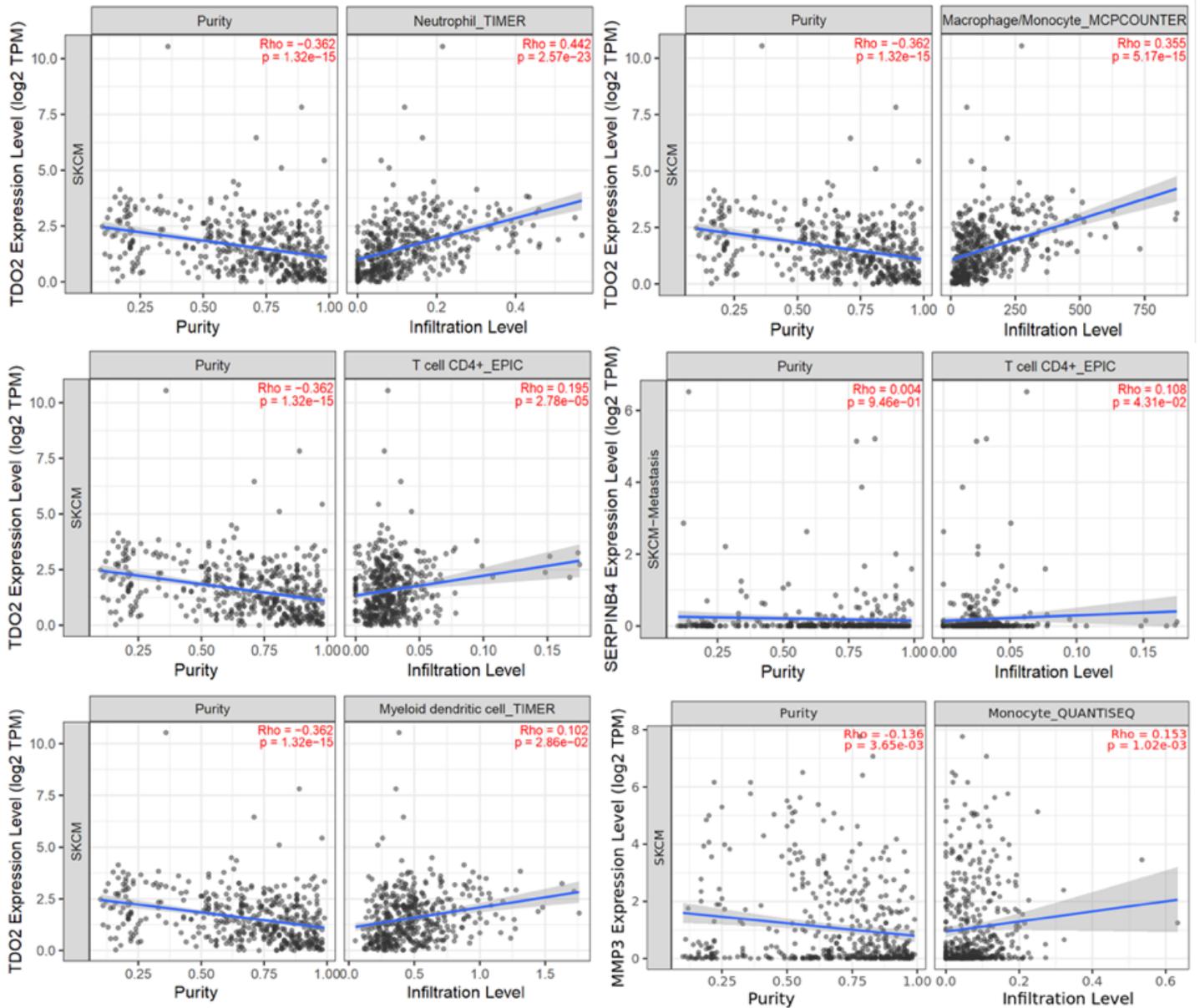


Figure 12

The correlation between TDO2, SERPINB4, and MMP3 genes and immune cells infiltration.

Supplementary Files

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

- [Tables.docx](#)